

OM protein - protein search, using sw model

Run on: January 26, 2004, 19:04:49 ; Search time 73 Seconds
(without alignments)
28.266 Million cell updates/sec

Title: US-09-766-412-30
Perfect score: 66
Sequence: 1 QPVLHLVALNTPL 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*
1: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1980.DAT.*
2: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1981.DAT.*
3: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1982.DAT.*
4: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1983.DAT.*
5: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1984.DAT.*
6: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1985.DAT.*
7: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1986.DAT.*
8: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1987.DAT.*
9: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1988.DAT.*
10: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1989.DAT.*
11: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1990.DAT.*
12: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1991.DAT.*
13: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1992.DAT.*
14: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1993.DAT.*
15: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1994.DAT.*
16: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1995.DAT.*
17: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1996.DAT.*
18: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1997.DAT.*
19: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1998.DAT.*
20: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA1999.DAT.*
21: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA2000.DAT.*
22: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA2001.DAT.*
23: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA2002.DAT.*
24: /SID1/gcgdata/Geneseq/Geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	100.0	13	22	Angiostatin-derive
2	66	100.0	13	22	Exemplary anti-ang
3	66	100.0	13	23	Antiangiogenic pep
4	66	100.0	20	18	N-terminal region
5	66	100.0	20	21	Mouse endostatin N
6	66	100.0	20	23	Endostatin N-termi
7	66	100.0	23	22	Murine endostatin
8	66	100.0	39	22	Antiangiogenic non
9	66	100.0	42	22	Murine sig-mEndo f

10	56	100.0	183	23	AAW49504	Mouse endostatin p
11	56	100.0	184	20	AAV18409	Endostatin protein
12	56	100.0	184	20	AAV08689	Murine endostatin
13	56	100.0	184	21	AAV70258	Murine angiogenesi
14	56	100.0	184	22	AAW49380	Murine endostatin
15	56	100.0	184	23	ABG31793	Human endostatin p
16	56	100.0	185	20	AAV06197	Anti-angiogenic en
17	56	100.0	191	21	AAW28398	Murine endostatin.
18	56	100.0	191	23	AAU77950	Amino acid sequenc
19	56	100.0	207	22	AAE02331	Murine endostatin
20	56	100.0	207	22	AAW71930	Murine endostatin
21	56	100.0	207	23	ABW79902	Mouse endostatin.
22	56	100.0	218	20	AAV08691	Murine gene therap
23	56	100.0	580	20	AAV08692	Murine gene therap
24	56	100.0	684	20	AAV25114	Mouse alpha1 (XVII
25	56	100.0	1288	18	AAW26328	Mouse alpha-1 coll
26	56	100.0	1288	20	AAW92297	Mouse alpha-1 (XVI
27	62	93.9	20	22	AAW49550	Human endostatin p
28	62	93.9	25	22	AAW49566	Human endostatin p
29	62	93.9	44	23	ABG31787	Human endostatin a
30	62	93.9	44	23	ABG31791	Human endostatin a
31	62	93.9	178	21	AAV94324	Alternate human en
32	62	93.9	178	22	AAU00900	Human Endostatin(T
33	62	93.9	179	22	AAU00901	Human Endostatin(T
34	62	93.9	180	22	AAU00899	Human Endostatin(T
35	62	93.9	181	22	AAV00898	Human Endostatin(T
36	62	93.9	182	21	AAW28399	Human endostatin.
37	62	93.9	182	21	AAV94323	Human endostatin p
38	62	93.9	182	21	AAV59622	Human endostatin p
39	62	93.9	182	22	AAU00897	Human Endostatin(T
40	62	93.9	182	23	AAU77951	Amino acid sequenc
41	62	93.9	183	20	AAV08693	Human endostatin p
42	62	93.9	183	20	AAV02113	SEQ ID 76 of W0991
43	62	93.9	183	21	AAW30493	Amino acid sequenc
44	62	93.9	183	21	AAW16451	Human endostatin p
45	62	93.9	183	21	AAV90771	Human angiogenesis

ALIGNMENTS

RESULT 1
AAB80862
ID AAB80862 standard; Peptide; 13 AA.
XX
AC AAB80862;
XX
DT 29-MAY-2001 (first entry)
XX
DE Angiostatin-derived peptide #19.
XX
KW Cytostatic; angiogenesis inhibitor; tumour metastasis; tumour growth.
XX
OS Unidentified.
XX
PN US6200954-B1.
XX
PD 13-MAR-2001.
XX
PF 30-AUG-1999; 99US-0385442.
XX
PR 04-SEP-1998; 98US-0099313.
XX
PA (UYSI-) UNIV SINGAPORE NAT.
XX
PI Ge R, Kini RM;
XX
DR WPI; 2001-234520/24.
XX
PT Peptides having potent anti-angiogenic activity, useful for the
XX treatment of tumors, comprise a portion of a FLT-1 protein -
PS Disclosure; Columns 15-16; 21pp; English.

XX The present sequence is a peptide which is effective in inhibiting
 CC undesirable angiogenesis. Angiogenesis is the process of new blood vessel
 CC formation from pre-existing vessels. Inappropriate angiogenesis is
 CC associated with various pathological conditions including solid tumour
 CC growth and metastasis. The present peptide can be used to prevent
 CC tumour metastasis or inhibit the growth of a primary tumour.

XX SQ Sequence 13 AA;

Query Match 100.0%; Score 66; DB 22; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.1e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTP 13
 |||||
 Db 1 QPVLHLVALNTP 13

RESULT 2

AAE74256

ID AAE74256 standard; peptide; 13 AA.

AC AAE74256;

DT 23-MAY-2001 (first entry)

DE Exemplary anti-angiogenic peptide #30.

XX Anti-angiogenesis; plasminogen; VEGF;

KW vascular endothelial growth factor; FLT-1; tumour; metastasis;
 KW cancer.

XX Homo sapiens.

XX WO200118030-A2.

XX 15-MAR-2001.

XX 01-SEP-2000; 2000WO-SG00131.

XX 03-SEP-1999; 99SG-0004310.

XX (UYSI-) UNIV SINGAPORE NAT.

XX Ge R, Kini RM;

XX WPI; 2001-257785/26.

XX Peptides comprising a portion of a protein selected from plasminogen,
 PT endostatin, VEGF, FLT-1 and KDR/FLK-1 are useful for treating primary
 PT tumor growth.

XX Claim 7; Page 18; 34pp; English.

XX The present invention relates to anti-angiogenesis peptides from
 CC a portion of a selected from plasminogen, endostatin, VEGF or FLT-1.
 CC The invention is used to prevent or treat primary tumour growth or
 CC metastasis or undesired angiogenesis.

XX SQ Sequence 13 AA;

Query Match 100.0%; Score 66; DB 22; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.1e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTP 13
 |||||
 Db 1 QPVLHLVALNTP 13

RESULT 3

ABG97539

ID ABG97539 standard; Peptide; 13 AA.

XX ABG97539;

DT 16-DEC-2002 (first entry)

DE Antiangiogenic peptide, Endo-2.

XX Angiogenesis; plasminogen; endostatin; Endo-; VEGF;

KW vascular endothelial growth factor; FLT-1; kinase insert domain; FLK-1;
 KW angiogenic inhibitor; Angio-; endothelial cell; proliferation;

KW tumour growth; blood vessel formation; migration; tubule structure;

KW rheumatoid arthritis; wound healing; tumour metastasis;

KW antiangiogenic therapy; mitogen; tyrosine kinase receptor.

XX Mammalia.

XX US2002103129-A1.

XX 01-AUG-2002.

XX 22-JAN-2001; 2001US-0766412.

XX 04-SEP-1998; 98US-099313P.

XX 30-AUG-1999; 99US-0385442.

XX (GERR/) GE R.

XX (KINI/) KINI R M.

XX Ge R, Kini RM;

XX WPI; 2002-749595/81.

XX Novel peptide for preventing/treating undesired angiogenesis, has a

PT portion of a protein e.g. plasminogen, endostatin, and potent

PT antiangiogenic activity and endothelial cell proliferation inhibition

XX activity

XX Claim 9; Page 10; 24pp; English.

XX The invention discloses a peptide comprising a portion of a protein e.g.
 CC plasminogen, endostatin (Endo-1-4), vascular endothelial growth factor
 CC (VEGF), FLT-1 or kinase insert domain containing receptor/FLK-1. These

CC peptides are able to act as potent angiogenic inhibitors (Angio-1-5)

CC inhibiting endothelial cell proliferation and retarding tumour growth.

CC Angiogenesis is the process of new blood vessel formation from

CC pre-existing vessel, involving endothelial cell proliferation, migration

CC and assembly into tubule structures. It plays an important role in normal

CC physiological functions such as embryonic development and wound healing.

CC Inappropriate angiogenesis is also associated with various pathological

CC conditions including tumour growth and metastasis, rheumatoid arthritis

CC and psoriasis. Anticancer therapy by inhibiting tumour angiogenesis is

CC called antiangiogenic therapy. VEGF is an endothelial specific mitogen

CC that functions through two high affinity tyrosine kinase receptors, FLT-1

CC and FLK-1 (not defined). Protein-protein interactions are crucial to many

CC physiological and pharmacological processes and the peptides disclosed

CC interfere with these interactions. The peptides are useful for preventing

CC or treating undesired angiogenesis and primary tumour growth or

CC metastasis. The sequences presented in ABG97510-ABG97559 are examples of

XX antiangiogenic peptides.

XX SQ Sequence 13 AA;

Query Match 100.0%; Score 66; DB 23; Length 13;
 Best Local Similarity 100.0%; Pred. No. 2.1e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTP 13
 |||||
 Db 1 QPVLHLVALNTP 13

RESULT 4

AAW16596
ID AAW16596 standard; peptide; 20 AA.

XX AC AAW16596;

XX DT 28-JAN-1998 (first entry)

XX DE N-terminal region of a novel endostatin.

XX KW Endostatin; angiogenesis; cancer; tumour; rheumatoid arthritis;
XX KW psoriasis; ocular; Osler-Webber Syndrome; myocardial; telangiectasia;
XX KW plaque neovascularisation; haemophilic joint; angiofibroma; inhibitor;
XX KW wound granulation; intestinal adhesion; atherosclerosis; scleroderma;
XX KW hypertrophic scar; cat scratch disease; Rochelle minialia quintosa;
XX KW Helicobacter pylori ulcer; birth control; collagen alpha type XVIII.

XX OS Mus sp.

XX PN WO9715666-A1.

XX PD 01-MAY-1997.

XX PF 23-OCT-1996; 96WO-US16925.

XX PR 22-OCT-1996; 96US-0740169.

XX PR 23-OCT-1996; 96US-0005835.

XX PR 02-AUG-1996; 96US-0023070.

XX PR 17-SEP-1996; 96US-0026263.

XX PA (CHIL-) CHILDRENS MEDICAL CENT.

XX PI Folkman MJ, O'Reilly MS;

XX WPI; 1997-259020/23.

XX PT Isolated endostatin - useful for treating, e.g. angiogenesis

XX PS dependent cancers

XX Claim 3; Page 56; 84pp; English.

XX CC A novel 20kDa endostatin that specifically inhibits endothelial
XX CC cell proliferation was isolated and is characterised by its preferred
XX CC N-terminal amino acid sequence. The present sequence represents the
XX CC first 20 of these amino acids. The N-terminal sequence corresponds
XX CC to an internal 20 amino acid fragment found in mouse collagen alpha type
XX CC XVIII satting at amino acid 1105 and ending at amino acid 1124. The
XX CC N-terminal amino acid sequence of the inhibitor also corresponds to an
XX CC internal 20 amino acid fragment found in human alpha 1 type XVIII
XX CC starting at amino acid 1132 and ending at amino acid 1151. Endostatin
XX CC can be isolated from murine hemangioendothelioma. The endostatin can be
XX CC used to treat angiogenesis related diseases, e.g. angiogenesis dependent
XX CC cancers, benign tumours, rheumatoid arthritis, psoriasis, ocular
XX CC angiogenesis diseases, Osler-Webber Syndrome, myocardial angiogenesis,
XX CC plaque neovascularisation, telangiectasia, haemophilic joints,
XX CC angiofibroma, wound granulation, intestinal adhesions, atherosclerosis,
XX CC scleroderma, hypertrophic scars, cat scratch disease (Rochelle minialia
XX CC quintosa) and Helicobacter pylori ulcers. The endostatin can also be used
XX CC to prevent embryo implantation, i.e. in birth control.

XX Sequence 20 AA;

Query Match 100.0%; Score 66; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13

Db 7 QPVLHLVALNTPL 19

RESULT 5

AAW94321

ID AAW94321 standard; peptide; 20 AA.

XX AC AAW94321;

XX DT 11-AUG-2000 (first entry)

XX DE Mouse endostatin N-terminus.

XX KW Mouse; endothelial cell proliferation inhibitor; collagen XVIII;
XX KW angiogenesis inhibitor; anti-tumour; cytostatic; antiproliferative;
XX KW vasotropic; dermatological; ophthalmological; vulnary;
XX KW antiarteriosclerotic; antidiabetic; haemostatic; contraceptive;
XX KW ocular angiogenic disease; atherosclerosis; scleroderma;
XX KW myocardial angiogenesis; telangiectasia; angiofibroma;
XX KW wound granulation.

XX OS Mus musculus.

XX PN WO200026368-A2.

XX PD 11-MAY-2000.

XX PF 01-NOV-1999; 99WO-US25605.

XX PR 30-OCT-1998; 98US-0106343.

XX PR 20-MAY-1999; 99US-0315689.

XX PA (CHIL-) CHILDRENS MEDICAL CENT.

XX PI O'Reilly MS, Folkman MJ;

XX WPI; 2000-365617/31.

XX PT Novel endostatin capable of inhibiting endothelial cell proliferation
XX PT and angiogenesis, useful for treating angiogenesis-dependent cancers
XX PT and as birth control agents

XX PS Disclosure; Page 2; 68pp; English.

XX CC The present sequence is the N-terminus of an endostatin, a potent and
XX CC specific protein inhibitor of endothelial cell proliferation in
XX CC vitro and angiogenesis in vivo. The sequence corresponds to a
XX CC C-terminal fragment of newly identified collagen type XVIII.
XX CC Recombinant mouse endostatin (20 mg/kg) was administered subcutaneously
XX CC to mice implanted with Lewis lung carcinomas. There was tumour mass
XX CC regression non-detectable levels after 12 days of therapy due to the
XX CC angiogenesis inhibitory activity of endostatin. Thus the protein is
XX CC useful for treatment of angiogenesis-dependent cancers. The
XX CC polynucleotide and polypeptide sequences of this endostatin are useful
XX CC for treating and diagnosis of tumours, ocular angiogenic diseases,
XX CC Osler-Webber syndrome, myocardial angiogenesis, plaque
XX CC neovascularisation, telangiectasia, haemophilic joints, angiofibroma
XX CC and wound granulation, for treatment of diseases related to excessive or
XX CC abnormal stimulation of endothelial cells e.g. intestinal adhesions,
XX CC atherosclerosis, scleroderma. The protein may also be useful as a birth
XX CC control agent by reducing or preventing uterine vascularisation. The
XX CC gene for endostatin may be isolated from cells or tissue that express
XX CC high levels of endostatin, eg. tumour cells, by generating cDNA from
XX CC mRNA using reverse transcriptase and then amplifying the DNA sequence.

XX Sequence 20 AA;

Query Match 100.0%; Score 65; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13

Db 7 QPVLHLVALNTPL 19

RESULT 6

AAW48921

ID AAM48821 standard; Peptide; 20 AA.
 XX
 AC AAM48821;
 XX
 DT 04-APR-2002 (first entry)
 XX
 DE Endostatin N-terminal peptide.
 XX
 KW Human; angiostatin; endostatin; angiogenesis; cancer; metastasis;
 KW psoriasis; scleroderma; Crohn's disease; corneal disease;
 KW retinopathy; arthritis; wound healing; Helicobacter pylori; peptic ulcer;
 KW gene therapy; angiostatin antagonist; endostatin antagonist;
 KW antiangiogenic; cytostatic; antiarthritic; antiinflammatory;
 KW cerebroprotective; antidiabetic; virucide; antipyretic; vulnery;
 KW gynaecological; cat scratch fever.
 XX
 OS Unidentified.
 XX
 PN WO200193897-A2.
 XX
 PD 13-DEC-2001.
 XX
 PF 04-JUN-2001; 2001WO-US17947.
 XX
 PR 02-JUN-2000; 2000US-209065P.
 PR 08-MAY-2001; 2001US-289387P.
 XX
 PA (ENTR-) ENTREMED INC.
 XX
 PI Sim XL, MacDonald NJ;
 XX
 DR WPI; 2002-130569/17.
 XX
 PR Regulating angiogenesis and treatment of angiogenesis-mediated
 PT diseases, e.g. hemangioma, tumors or cancer, by administering a
 PT tropomyosin binding compound or actin disrupting compound -
 XX
 PS Disclosure; Page 13; 95pp; English.
 XX
 CC The present invention relates to methods of regulating angiogenesis in an
 CC individual by administering an angiogenesis regulating composition
 CC comprising a tropomyosin binding compound or an actin disrupting
 CC compound. The compositions are useful for treating diseases and processes
 CC mediated by angiogenesis including haemangioma, solid tumours, blood
 CC borne tumours, leukaemia, metastasis, Crohn's disease, coronary or
 CC cerebral collateral, arthritis, diabetic neovascularisation, macular
 CC degeneration, wound healing, Helicobacter related diseases, ovulation,
 CC menstruation, and cat scratch fever. The present sequence is a peptide
 CC described in the exemplification of the invention.
 XX
 SQ Sequence 20 AA;
 Query Match 100.0%; Score 66; DB 23; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3.4e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QPVLHLVALNTPL 13
 DB 7 QPVLHLVALNTPL 19
 RESULT 7
 AAB49806
 ID AAB49806 standard; Peptide; 23 AA.
 XX
 AC AAB49806;
 XX
 DT 02-MAR-2001 (first entry)
 XX
 DE Murine endostatin peptide fragment SEQ ID NO: 19.
 XX
 KW Endostatin; antiangiogenic; angiogenesis; human; mouse; chicken;
 KW cancer; inflammation; angiogenesis-dependent disease.

XX Mus musculus.
 OS
 PN WO200067771-A1.
 XX
 PD 16-NOV-2000.
 XX
 PF 02-MAY-2000; 2000WO-US12063.
 XX
 PR 06-MAY-1999; 99US-0132907.
 PR 14-JUL-1999; 99US-0353333.
 XX
 PA (BURN-) BURNHAM INST.
 XX
 PI Vuori K;
 XX
 DR WPI; 2001-040937/05.
 XX
 PT Endostatin peptide comprising at least four endostatin amino acid
 PT residues are e.g. angiogenesis inhibitors for treating cancer and
 PT diabetic retinopathy -
 XX
 PS Example 4; Page 124; 146pp; English.
 XX
 CC The present invention provides endostatin peptides which can be used in
 CC the modulation of angiogenesis. This is useful in the treatment of
 CC cancers, inflammation, rheumatoid arthritis, chronic articular
 CC rheumatism, psoriasis, disorders associated with inopportune invasion of
 CC vessels such as diabetic retinopathy, neovascular glaucoma, retinopathy
 CC of prematurity, macular degeneration, corneal graft rejection,
 CC retrolental fibroplasia, rubeosis, capillary proliferation in,
 CC atherosclerotic plaques and osteoporosis. Other angiogenesis-dependent
 CC diseases include Osler-Webber syndrome, myocardial angiogenesis, plaque
 CC neovascularisation, telangiectasia, haemophilic joints and wound
 CC granulation. In addition, the peptides can be used as birth control
 CC agents.
 XX
 SQ Sequence 23 AA;
 Query Match 100.0%; Score 66; DB 22; Length 23;
 Best Local Similarity 100.0%; Pred. No. 3.9e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QPVLHLVALNTPL 13
 DB 7 QPVLHLVALNTPL 19
 RESULT 8
 AAB35584
 ID AAB35584 standard; peptide; 39 AA.
 XX
 AC AAB35584;
 XX
 DT 14-FEB-2001 (first entry)
 XX
 DE Antiangiogenic nonatriacenta peptide I.
 XX
 KW Antiangiogenic; angiogenesis; cancer; endostatin.
 XX
 OS Synthetic.
 XX
 PN WO200063249-A1.
 XX
 PD 26-OCT-2000.
 XX
 PF 11-APR-2000; 2000WO-EF03236.
 XX
 PR 15-APR-1999; 99IT-MI00777.
 XX
 PA (UYMI-) UNIV MILANO.
 PA (UYFI-) UNIV FIRENZE.
 XX

PI Chillemi F, Francescato P, Ziche M;
DR WPI; 2001-007005/01.
XX
XX Polypeptides derived from endostatin exhibiting antiangiogenic activity
PT useful for treatment of angiogenesis-dependent tumours
XX
XX Claim 2; Page 18; 28pp; English.
PS
XX The present invention describes a number of peptides derived from
CC endostatin which exhibit antiangiogenic activity. These may be used in
CC the treatment of cancer. The present sequence is one of the peptides of
CC the invention.
XX
XX SQ Sequence 39 AA;
Query Match 100.0%; Score 66; DB 22; Length 39;
Best Local Similarity 100.0%; Pred. No. 7.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19
RESULT 9
AAB71931
ID AAB71931 standard; Protein; 42 AA.
XX
XX AAB71931;
AC
XX
XX 10-MAY-2001 (first entry)
DT
XX
XX Murine sig-mEndo fusion protein N-terminal sequence.
DE
XX
XX Mouse; endostatin; antitumor; cytostatic; antiarthritic; antipsoriatic;
KW antidiabetic; ophthalmological; gene therapy; angiogenic inhibitor;
KW adenoviral vector; diabetic retinopathy; cardiovascular disease;
KW arthritis; psoriasis; cerebral oedema; intravascular coagulopathy;
KW lymphoma; leukaemia; sig-mEndo; fusion protein.
XX
XX OS Mus sp.
XX
XX WO200112830-A1.
FN
XX
XX 22-FEB-2001.
PD
XX
XX 11-AUG-2000; 2000WO-EP07865.
PF
XX
XX 13-AUG-1999; 99US-0373938.
PR
XX
XX (NOVS) NOVARTIS AG.
PA
XX (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
XX
PI Hallenbeck PL, Chen CT;
XX
XX WPI; 2001-202871/20.
DR
XX
XX Adenoviral vector for treating tumors and disorders associated with
PT angiogenesis, such as cancer, arthritis, and psoriasis, comprises a DNA
PT sequence encoding an angiogenic inhibitor, particularly endostatin -
XX
XX Example 1; Fig 5; 59pp; English.
PS
XX The present peptide was analysed in a study of the expression and
CC analysis of sig-mEndo, which comprises murine endostatin attached to the
CC murine Ig-kappa signal sequence. sig-mEndo was introduced into an
CC adenoviral vector. The adenoviral vector is useful for expressing
CC endostatin in a mammalian cell such as an A549 or Hep3B cell. It is
CC useful for treating other diseases and disorders associated with
CC angiogenesis, such as neovascular diseases of the eye, including diabetic
CC retinopathy, cardiovascular disease, arthritis, psoriasis, cerebral
CC oedema and intravascular coagulopathy (Kasabach-Merritt syndrome). The

CC vector inhibits, prevents or destroys the growth of tumours by
CC preventing the formation of blood vessels in tumours, such as lymphoma
CC and leukaemia.
XX
XX SQ Sequence 42 AA;
Query Match 100.0%; Score 66; DB 22; Length 42;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QPVLHLVALNTPL 13
Db 30 QPVLHLVALNTPL 42
RESULT 10
AAM49504
ID AAM49504 standard; Protein; 183 AA.
XX
XX AAM49504;
AC
XX
XX 07-MAY-2002 (first entry)
DT
XX
XX Mouse endostatin protein.
DE
XX
XX Endostatin; murine; proliferation; blood vessel endothelium;
KW regeneration; tumour; blood vessel; treatment; amplification.
KW
XX
XX OS Mus sp.
XX
XX Key Location/Qualifiers
FH Misc-difference 181
FT FT /label= Ser, Phe
FT
XX
XX CN1177005-A.
FN
XX
XX 25-MAR-1998.
PD
XX
XX 10-SEP-1997; 97CN-0107112.
PF
XX
XX 10-SEP-1997; 97CN-0107112.
PR
XX
XX (XUGG/) XU G.
PA
XX
XX Xu G, Ren M, Xu L;
PI
XX
XX WPI; 2002-106746/15.
DR
XX
XX Gene clone of inhibitory factor for hyperplasia of inner blood vessel
PT cells in human body's real tumor, and its use in anti-tumor blood
PT vessel regeneration -
PT
XX
XX Disclosure; Page 4 (Disclosure); 6pp; Chinese.
PS
XX This invention describes a novel preparation which inhibits the
CC proliferation of blood vessel endothelium and prevents the regeneration
CC activity of tumour blood vessels. The preparation can also be used as a
CC biological preparation in the treatment of tumours. This sequence
CC represents the murine endostatin protein described in the invention.
XX
XX SQ Sequence 183 AA;
Query Match 100.0%; Score 66; DB 23; Length 183;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19
RESULT 11
AAV18409

ID XX AAY18409 standard; Protein; 184 AA.
AC XX AAY18409;
XX XX
DT 24-AUG-1999 (first entry)
XX XX Endostatin protein sequence.
DE XX
XX XX
XX EM1; anti-angiogenic peptide; endostatin; angiogenesis-dependent cancer;
KW benign tumour; rheumatoid arthritis; psoriasis; ocular angiogenesis;
KW Osler-Webber Syndrome; myocardial angiogenesis; angiofibroma; cancer;
KW plaque neovascularisation; telangiectasia; atherosclerosis; scleroderma;
KW dialysis graft vascular access stenosis; renal cancer; therapy.
XX XX
OS Mus sp.
XX WO9929855-A1.
XX 17-JUN-1999.
XX 08-DEC-1998; 98WO-US26057.
XX 16-NOV-1998; 98US-0108536.
PR 08-DEC-1997; 97US-0067888.
PR 22-APR-1998; 98US-0082663.
XX (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
XX Sukhatme VP;
XX WPI; 1999-385604/32.
DR N-PSDB; AAX79949.
XX Mutant endostatin having anti-angiogenic activity
PT Claim 31; Fig 2; 105pp; English.
PS This sequence is the mouse endostatin. The invention relates to a
XX the mutant endostatin (EM1), which has anti-angiogenic activity, and is
CC designated EM1. Compositions comprising EM1 or fusion proteins comprising
CC EM1, are useful for treating diseases characterised by angiogenic
CC activity, such as angiogenesis-dependent cancers, benign tumours,
CC rheumatoid arthritis, psoriasis, ocular angiogenesis, Osler-Webber
CC Syndrome, myocardial angiogenesis, plaque neovascularisation,
CC telangiectasia, haemophilic joints, angiofibroma, wound granulation,
CC intestinal adhesions, atherosclerosis, scleroderma, hypertrophic scars,
CC cat scratch disease, Helicobacter pylori ulcers, dialysis graft vascular
CC access stenosis, contraception and obesity. In particular, the diseases
CC treatable by EM1 comprise cancer, especially renal cancer. The methods
CC provide a means for introducing EM1 into mammalian cells via gene
CC therapy, for production of EM1 via recombinant means, as well as
CC recombinant production of the EM1 protein. EM1 performs as well or better
CC than whole endostatin. Use of EM1 is advantageous for treatment of
CC angiogenic diseases in that increasingly smaller peptides are more potent
CC on a weight basis, and may be able to better penetrate tissues.
XX XX
SQ Sequence 184 AA;
Query Match 100.0%; Score 66; DB 20; Length 184;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QPVLHLVALNTPL 13
DB 7 QPVLHLVALNTPL 19
RESULT 12
AAY08689
ID AAY08689 standard; Protein; 184 AA.
XX XX
AC AAY08689;
XX XX

DT 10-AUG-1999 (first entry)
XX XX Murine endostatin protein fragment.
DE XX
XX Plasmidogen; murine; angiostatin; endostatin; gene therapy; vector;
KW anti-angiogenic; attenuation; cytostatic; anti-diabetic; ophthalmology;
XX tumour growth; solid tumour; diabetic retinopathy; retina.
XX OS Mus sp.
XX WO9926480-A1.
XX 03-JUN-1999.
XX 20-NOV-1998; 98WO-US24950.
XX 20-NOV-1997; 97US-0975424.
PR (GENE-) GENETIX PHARM INC.
XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
PA Rachelot T, Leboulch P, Pawliuk RJ;
XX WPI; 1999-357696/30.
DR N-PSDB; AAX77715.
XX Anti-angiogenic gene therapy vectors
PT Disclosure; Fig 6; 83pp; English.
XX This invention describes a novel viral gene therapy vector comprising a
CC nucleic acid molecule encoding an anti-angiogenic polypeptide chosen
CC from human or murine angiostatin, human or murine endostatin and
CC angiogenesis-inhibiting fusions and fragments, where the viral vector is
CC sufficiently attenuated for use in human gene therapy. The products of
CC the invention have anti-angiogenic, cytostatic, anti-diabetic and
CC ophthalmological activity. The vector is used in gene therapy for
CC inhibiting tumour growth in humans harbouring a solid tumour. The vector
CC expresses an anti-angiogenic polypeptide. An additional use comprises
CC treatment of diabetic retinopathy, where the anti-angiogenic polypeptide
CC inhibits angiogenesis in the vicinity of the retina. The vector is
CC administered to cells ex vivo and then administered to the patient.
XX XX
SQ Sequence 184 AA;
Query Match 100.0%; Score 66; DB 20; Length 184;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QPVLHLVALNTPL 13
DB 7 QPVLHLVALNTPL 19
RESULT 13
AAY70258
ID AAY70258 standard; Protein; 184 AA.
XX XX
AC AAY70258;
XX 06-JUN-2000 (first entry)
XX Murine angiogenesis inhibitor, endostatin.
XX Murine; immunoglobulin Fc fragment; endostatin; immunofusin;
KW angiogenesis; inhibitor; cytostatic; antirheumatoid; antiarthritic;
KW antidiabetic; antidiabetic; ophthalmological; immunosuppressant;
KW vasotropic; vulvar; treatment; antiarteriosclerosis; tumour;
KW metastasis; atherosclerosis; psoriasis; rheumatoid arthritis;
KW ocular angiogenic disease; diabetic retinopathy; macular degeneration;
KW myocardial angiogenesis; plaque neovascularisation; telangiectasia;
KW wound granulation; keloid scar; gene therapy.
XX XX

OS Mus musculus.
 XX WO200011033-A2.
 XX
 XX 02-MAR-2000.
 XX
 XX 25-AUG-1999; 99WO-US19329.
 XX
 XX 25-AUG-1998; 98US-0097883.
 XX
 XX (LEXI-) LEXINGEN PHARM CORP.
 XX
 XX Lo K, Li Y, Gillies SD;
 XX
 XX WPI; 2000-237616/20.
 XX
 XX N-PSDB; AAZ51299.
 XX
 XX Novel fusion protein of angiostatin or endostatin and an immunoglobulin
 PT FC region, useful for treating conditions mediated by angiogenesis,
 PT such as rheumatoid arthritis, tumors and macular degeneration -
 XX
 XX Example 5; Pages 48-49; 68pp; English.
 XX
 XX The patent discloses a DNA molecule encoding a fusion protein comprising
 CC a signal sequence, an immunoglobulin Fc region, and an angiogenesis
 CC inhibitor selected from angiostatin, endostatin, and a plasminogen fragment
 CC having angiostatin activity, a collagen XVIII fragment having endostatin
 CC activity, or combinations of them. The fusion protein (immunofusin) is
 CC used to inhibit angiogenesis and to treat diseases or conditions mediated
 CC by angiogenesis. Conditions that may be treated include solid tumours,
 CC blood born tumours, tumour metastasis, benign tumours including
 CC haemangiomas, acoustic neuromas, neurofibromas, trachomas and pyrogenic
 CC granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases
 CC e.g. diabetic retinopathy, retinopathy of prematurity, macular
 CC degeneration, corneal graft rejection, neovascular glaucoma, retrolental
 CC fibroplasia, rubeosis and Osler-Webber syndrome; myocardial angiogenesis,
 CC plaque neovascularisation, telangiectasia, haemophilic joints,
 CC angiofibroma, wound granulation, and excessive or abnormal stimulation of
 CC endothelial cells, intestinal cells, atherosclerosis, scleroderma and
 CC hypertrophic scars, i.e. keloid scars. The DNA constructs may be used
 CC in gene therapy. The present sequence is a murine
 CC endostatin used in the construction of immunofusin containing murine
 CC immunoglobulin Fc fragment.
 XX
 XX Sequence 184 AA;
 Query Match 100.0%; Score 66; DB 21; Length 184;
 Best Local Similarity 100.0%; Pred. No. 0.00041;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QPVLHLVALNTPL 13
 DB 7 QPVLHLVALNTPL 19
 RESULT 14
 AAB49380
 ID AAB49380 standard; Protein; 184 AA.
 XX
 XX AAB49380;
 XX
 XX 02-MAR-2001 (first entry)
 DT
 XX Murine endostatin SEQ ID NO: 4.
 DE
 XX Endostatin; antiangiogenic; angiogenesis; human; mouse; chicken;
 XX cancer; inflammation; angiogenesis-dependent disease.
 XX
 XX Mus musculus.
 OS
 XX WO200067771-A1.
 XX
 XX 16-NOV-2000.

XX 02-MAY-2000; 2000WO-US12063.
 XX
 XX 06-MAY-1999; 99US-0132907.
 XX
 XX 14-JUL-1999; 99US-0353333.
 XX
 XX (BURN-) BURNHAM INST.
 XX
 XX Vuori K;
 XX
 XX WPI; 2001-040937/05.
 XX
 XX N-PSDB; AAC68290.
 XX
 XX Endostatin peptide comprising at least four endostatin amino acid
 PT residues are e.g. angiogenesis inhibitors for treating cancer and
 PT diabetic retinopathy -
 XX
 XX Disclosure; Fig 1; 146pp; English.
 XX
 XX The present invention provides endostatin peptides which can be used in
 CC the modulation of angiogenesis. This is useful in the treatment of
 CC cancers, inflammation, rheumatoid arthritis, chronic articular
 CC rheumatism, psoriasis, disorders associated with inopportune invasion of
 CC vessels such as diabetic retinopathy, neovascular glaucoma, retinopathy
 CC of prematurity, macular degeneration, corneal graft rejection,
 CC retrolental fibroplasia, rubeosis, capillary proliferation in
 CC atherosclerotic plaques, and osteoporosis. Other angiogenesis-dependent
 CC diseases include Osler-Webber syndrome, myocardial angiogenesis, plaque
 CC neovascularisation, telangiectasia, haemophilic joints and wound
 CC granulation. In addition, the peptides can be used as birth control
 CC agents.
 XX
 XX Sequence 184 AA;
 Query Match 100.0%; Score 66; DB 22; Length 184;
 Best Local Similarity 100.0%; Pred. No. 0.00041;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QPVLHLVALNTPL 13
 DB 7 QPVLHLVALNTPL 19
 RESULT 15
 ABG31793
 ID ABG31793 standard; Protein; 184 AA.
 XX
 XX ABG31793;
 AC
 XX 05-DEC-2002 (first entry)
 DT
 XX Human endostatin polypeptide.
 DE
 XX Human; endostatin; tumour; cancer; metastasis; cytostatic;
 XX antiangiogenic.
 XX
 XX Homo sapiens.
 OS
 XX WO200269457-A2.
 XX
 XX 06-SEP-2002.
 XX
 XX 27-FEB-2002; 2002WO-IT00119.
 XX
 XX 27-FEB-2001; 2001IT-MI00394.
 XX
 XX (UYMI-) UNIV MILANO.
 XX
 XX Chillemi F, Vicentinie LMT, Francescato P;
 XX
 XX WPI; 2002-698655/75.
 XX
 XX New peptide useful for the preparation of medicaments with

PT antiangiogenic activity that may be used in treating tumours or
PT metastases, comprises a sequence corresponding to fragments of human
PT endostatin -

XX
XX Disclosure; Fig 1; 24pp; English.

XX The invention relates to peptide comprising 20-50 amino acids with
CC sequences corresponding to the human endostatin polypeptide sequence, its
CC salt or non-toxic derivative. The peptides are useful in the preparation
CC of medicaments with antiangiogenic activity which may be useful in
CC treating tumours or metastases. This sequence represents a human
CC endostatin polypeptide.

XX
SQ Sequence 184 AA;

Query Match 100.0%; Score 66; DB 23; Length 184;
Best Local Similarity 100.0%; Pred. NO. 0.00041;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
| | | | | | | | | |
Db 7 QPVLHLVALNTPL 19

Search completed: January 26, 2004, 19:06:22
Job time : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2004, 19:04:52 ; Search time 28 Seconds
(without alignments)
19.644 Million cell updates/sec

Title: US-09-766-412-30

Perfect score: 66

Sequence: 1 QPVLHLVALNTPL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	100.0	13	3	US-09-385-442-30
2	66	100.0	20	2	US-08-740-168A-1
3	66	100.0	20	3	US-09-349-429-1
4	66	100.0	20	4	US-09-315-689-1
5	66	100.0	20	4	US-09-174-282-1
6	66	100.0	185	3	US-08-985-526-36
7	66	100.0	191	4	US-09-561-500-13
8	66	100.0	191	4	US-09-561-108-13
9	66	100.0	191	4	US-09-561-526-13
10	66	100.0	191	4	US-09-561-499-13
11	66	100.0	195	1	US-08-159-784-2
12	62	93.9	178	4	US-09-315-689-5
13	62	93.9	182	4	US-09-561-500-14
14	62	93.9	182	4	US-09-561-108-14
15	62	93.9	182	4	US-09-315-689-3
16	62	93.9	182	4	US-09-561-526-14
17	62	93.9	182	4	US-09-561-499-14
18	62	93.9	183	3	US-09-206-059-2
19	44	66.7	191	1	US-08-159-784-3
20	41	62.1	955	4	US-09-252-991A-18882
21	40	60.6	413	4	US-09-724-623-87
22	38	57.6	2311	4	US-08-934-386-9
23	37	56.1	187	1	US-08-471-058-17
24	37	56.1	187	3	US-08-471-057-17
25	37	56.1	187	4	US-08-470-865-17
26	37	56.1	191	1	US-08-607-269-28
27	37	56.1	191	5	PCT-US95-04600-28

28 37 56.1 2475 3 US-09-413-814-48 Sequence 48, Appl
29 36 54.5 131 4 US-09-107-532A-4649 Sequence 4649, Ap
30 36 54.5 334 4 US-09-252-991A-29546 Sequence 29546, A
31 36 54.5 342 3 US-09-330-611-4 Sequence 4, Appl
32 36 54.5 511 3 US-08-991-677-4 Sequence 4, Appl
33 35 53.0 84 4 US-09-300-008B-60 Sequence 60, Appl
34 35 53.0 108 4 US-09-732-210-205 Sequence 205, App
35 35 53.0 222 4 US-09-134-001C-4748 Sequence 4748, Ap
36 35 53.0 287 4 US-09-252-991A-19928 Sequence 19928, A
37 35 53.0 378 4 US-09-186-489-2 Sequence 2, Appl
38 35 53.0 445 4 US-10-043-665B-2 Sequence 26149, A
39 35 53.0 472 4 US-09-252-991A-17229 Sequence 17229, A
40 35 53.0 672 4 US-09-252-991A-19440 Sequence 19440, A
41 35 53.0 951 4 US-09-252-991A-19440 Sequence 5, Appl
42 35 53.0 1090 3 US-09-085-199B-5 Sequence 742, App
43 34 51.5 172 4 US-09-198-452A-742 Sequence 1677, Ap
44 34 51.5 250 4 US-09-702-705-1677 Sequence 1677, Ap
45 34 51.5 250 4 US-09-736-457-1677 Sequence 1677, Ap

ALIGNMENTS

RESULT 1
US-09-385-442-30
; Sequence 30, Application US/09385442
; Patent No. 6200954
; GENERAL INFORMATION:
; APPLICANT: Ge, Ruowen
; APPLICANT: Kini, R. Manjunatha
; TITLE OF INVENTION: Small Peptides Having Potent Anti-Angiogenic Activity
; FILE REFERENCE: 1781-170P
; CURRENT APPLICATION NUMBER: US/09/385,442
; CURRENT FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 60/099,313
; EARLIER FILING DATE: 1999-09-04
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 13
; TYPE: PRT
; ORGANISM: mammalian
; FEATURE:
; OTHER INFORMATION: Endo-2
US-09-385-442-30

Query Match 100.0%; Score 66; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 3.6e-05; Indels 0; Gaps 0;
Matches 13; Conservative 0; Mismatches 0;

QY 1 QPVLHLVALNTPL 13
| | | | | | | | | | | | | | |
Db 1 QPVLHLVALNTPL 13

RESULT 2
US-08-740-168A-1
; Sequence 1, Application US/08740168A
; Patent No. 5854205
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions
; TITLE OF INVENTION: and Methods
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/740,168A
FILING DATE: 22-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0223
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Murine
TISSUE TYPE: Collagen
US-08-740-168A-1

Query Match 100.0%; Score 66; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 7 QPVLHLVALNTPL 19

RESULT 3
US-09-349-429-1
Sequence 1, Application US/09349429
Patent No. 6174861
GENERAL INFORMATION:
APPLICANT: O'Reilly, Michael
APPLICANT: Folkman, M. Judah
TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions
TITLE OF INVENTION: and Methods
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew, LLP
STREET: 191 Peachtree, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/349,429
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/740,168
FILING DATE: 22-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Warren, William L.
REGISTRATION NUMBER: 36,714
REFERENCE/DOCKET NUMBER: 05213-0223
TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-818-3700
TELEFAX: 404-818-3799
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Murine
TISSUE TYPE: Collagen
US-09-349-429-1

Query Match 100.0%; Score 66; DB 3; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 7 QPVLHLVALNTPL 19

RESULT 4
US-09-315-689-1
Sequence 1, Application US/09315689
Patent No. 6346510
GENERAL INFORMATION:
APPLICANT: Folkman, Judah
APPLICANT: O'Reilly, Michael
TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
FILE REFERENCE: 05213-0229
CURRENT APPLICATION NUMBER: US/09/315,689
CURRENT FILING DATE: 1999-05-20
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 20
TYPE: PRT
ORGANISM: murine
US-09-315-689-1

Query Match 100.0%; Score 66; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 7 QPVLHLVALNTPL 19

RESULT 5
US-09-174-282-1
Sequence 1, Application US/09174282
Patent No. 6544758
GENERAL INFORMATION:
APPLICANT: O'Reilly, Michael
APPLICANT: Folkman, M. Judah
TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions
TITLE OF INVENTION: and Methods
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jones & Askew, LLP
STREET: 191 Peachtree, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/174,282
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/740,168
; FILING DATE: 22-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0223
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; TISSUE TYPE: Collagen
; US-09-174-282-1

Query Match 100.0%; Score 66; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.8e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
| | | | | | | | | | | | | | | | | |
DB 7 QPVLHLVALNTPL 19

RESULT 6
US-08-985-526-36
; Sequence 36, Application US/08985526
; Patent No. 6080728
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA
; TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; TITLE OF INVENTION: THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,526
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/608,845
; FILING DATE: 16-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: McMorow Jr., Robert G
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (302) 658-9141
; TELEFAX: (302) 658-5613
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 185 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-985-526-36

Query Match 100.0%; Score 66; DB 3; Length 185;
Best Local Similarity 100.0%; Pred. No. 0.00064;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
| | | | | | | | | | | | | | | | | |
DB 8 QPVLHLVALNTPL 20

RESULT 7
US-09-561-500-13
; Sequence 13, Application US/09561500
; Patent No. 6342219
; GENERAL INFORMATION:
; APPLICANT: Phillip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002500
; CURRENT APPLICATION NUMBER: US/09/561,500
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; US-09-561-500-13

Query Match 100.0%; Score 66; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
| | | | | | | | | | | | | | | | | |
DB 14 QPVLHLVALNTPL 26

RESULT 8
US-09-561-108-13
; Sequence 13, Application US/09561108
; Patent No. 6342221
; GENERAL INFORMATION:
; APPLICANT: Phillip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/561,108
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; US-09-561-108-13

Query Match 100.0%; Score 66; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 14 QPVLHLVALNTPL 26

RESULT 9
US-09-561-526-13
; Sequence 13, Application US/09561526
; Patent No. 6416758
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE KITS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002586
; CURRENT APPLICATION NUMBER: US/09/561,526
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-526-13

Query Match 100.0%; Score 66; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 14 QPVLHLVALNTPL 26

RESULT 10
US-09-561-499-13
; Sequence 13, Application US/09561499
; Patent No. 6524583
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002582
; CURRENT APPLICATION NUMBER: US/09/561,499
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131,432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
US-09-561-499-13

Query Match 100.0%; Score 66; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 0.00067;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 14 QPVLHLVALNTPL 26

RESULT 11
US-08-159-784-2
; Sequence 2, Application US/08159784
; Patent No. 5643783
; GENERAL INFORMATION:
; APPLICANT: Bjorn R. Olsen
; TITLE OF INVENTION: NOVEL COLLAGEN AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/159,784
; FILING DATE: December 1, 1993
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: John F. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00246/170001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 195
; TYPE: amino acid
; STRANDEDNESS: N/A
; TOPOLOGY: N/A
US-08-159-784-2

Query Match 100.0%; Score 66; DB 1; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.00068;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 18 QPVLHLVALNTPL 30

RESULT 12
US-09-315-689-5
; Sequence 5, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Polkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315,689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 178
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-315-689-5

Query Match 93.9%; Score 62; DB 4; Length 178;
Best Local Similarity 92.3%; Pred. No. 0.0029;

Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 QPVLHLVALNTPL 13
Db 3 QPVLHLVALNSPL 15

RESULT 13
US-09-561-500-14
; Sequence 14, Application US/09561500
; Patent No. 6342219
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002500
; CURRENT APPLICATION NUMBER: US/09/561.500
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131.432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-561-500-14

Query Match 93.9%; Score 62; DB 4; Length 182;
Best Local Similarity 92.3%; Pred. No. 0.003;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNSPL 19

RESULT 14
US-09-561-108-14
; Sequence 14, Application US/09561108
; Patent No. 6342221
; GENERAL INFORMATION:
; APPLICANT: Philip E. Thorpe
; APPLICANT: Rolf A. Brekken
; TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY INHIBITING VEGF
; FILE REFERENCE: 4001.002584
; CURRENT APPLICATION NUMBER: US/09/561.108
; CURRENT FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/131.432
; PRIOR FILING DATE: 1999-04-28
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 14
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
; OTHER INFORMATION: PEPTIDE
US-09-561-108-14

Query Match 93.9%; Score 62; DB 4; Length 182;
Best Local Similarity 92.3%; Pred. No. 0.003;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNSPL 19

RESULT 15
US-09-315-689-3
; Sequence 3, Application US/09315689
; Patent No. 6346510
; GENERAL INFORMATION:
; APPLICANT: Folkman, Judah
; APPLICANT: O'Reilly, Michael
; TITLE OF INVENTION: Therapeutic Antiangiogenic Endostatin Compositions
; FILE REFERENCE: 05213-0229
; CURRENT APPLICATION NUMBER: US/09/315.689
; CURRENT FILING DATE: 1999-05-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-315-689-3

Query Match 93.9%; Score 62; DB 4; Length 182;
Best Local Similarity 92.3%; Pred. No. 0.003;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNSPL 19

Search completed: January 26, 2004, 19:10:12
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2004, 19:06:31 ; Search time 367 Seconds
(without alignments)
7,323 Million cell updates/sec

Title: US-09-766-412-30

Perfect score: 66

Sequence: 1 QPVLHLVALNTPL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 777136 seqs, 206736638 residues

Total number of hits satisfying chosen parameters: 777136

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	100.0	13	10	US-09-766-412-30
2	66	100.0	20	9	US-09-873-676-115
3	66	100.0	20	10	US-09-405-499-1
4	66	100.0	20	10	US-09-174-516-1
5	66	100.0	20	10	US-09-154-302-1
6	66	100.0	20	12	US-10-351-284-1
7	66	100.0	20	15	US-10-131-241-43
8	66	100.0	20	15	US-10-232-316-1
9	66	100.0	20	15	US-10-042-347-1
10	66	100.0	184	12	US-10-292-418-18
11	66	100.0	184	15	US-10-131-241-46
12	66	100.0	185	14	US-10-036-869-36
13	66	100.0	191	10	US-09-998-831-13
14	66	100.0	191	12	US-10-373-561-13
15	66	100.0	207	12	US-10-422-934-71

ALIGNMENTS

RESULT 1

US-09-766-412-30
; Sequence 30, Application US/09766412
; Patent No. US20020103129A1
; GENERAL INFORMATION:
; APPLICANT: GE. Ruwen et al.
; TITLE OF INVENTION: SMALL PEPTIDES HAVING ANTI-ANGIOGENIC AND ENDOTHELIAL CELL INHIBIT
; TITLE OF INVENTION: ACTIVITY
; FILE REFERENCE: 1781-0215P
; CURRENT APPLICATION NUMBER: US/09/766,412
; CURRENT FILING DATE: 2001-01-11
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 30
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Mammalian
; NAME/KEY: misc feature
; OTHER INFORMATION: Endo-2
US-09-766-412-30

Query Match 100.0%; Score 66; DB 10; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13

Db 1 QPVLHLVALNTPL 13

RESULT 2

US-09-873-676-115
; Sequence 115, Application US/09873676
; Patent No. US20020077289A1
; GENERAL INFORMATION:
; APPLICANT: MacDonald, Nicholas J.
; APPLICANT: Sim, Kim L.

16	66	100.0	207	14	US-10-080-797-3	Sequence 3, Appli
17	62	93.9	178	15	US-10-131-241-60	Sequence 60, Appl
18	62	93.9	178	15	US-10-042-347-5	Sequence 5, Appli
19	62	93.9	179	15	US-10-131-241-57	Sequence 57, Appl
20	62	93.9	180	15	US-10-131-241-47	Sequence 47, Appl
21	62	93.9	180	15	US-10-131-241-56	Sequence 56, Appl
22	62	93.9	181	15	US-10-131-241-55	Sequence 55, Appl
23	62	93.9	182	10	US-09-998-831-14	Sequence 14, Appl
24	62	93.9	182	12	US-10-373-561-14	Sequence 54, Appl
25	62	93.9	182	15	US-10-131-241-54	Sequence 54, Appl
26	62	93.9	182	15	US-10-042-347-3	Sequence 3, Appli
27	62	93.9	183	9	US-09-873-676-2	Sequence 2, Appli
28	62	93.9	183	12	US-10-292-418-4	Sequence 4, Appli
29	62	93.9	183	14	US-10-080-797-1	Sequence 1, Appli
30	62	93.9	183	15	US-10-131-241-52	Sequence 52, Appl
31	62	93.9	682	12	US-10-264-049-3010	Sequence 3010, Ap
32	62	93.9	684	11	US-09-961-403-5	Sequence 5, Appli
33	62	93.9	1516	12	US-10-431-642-3	Sequence 3, Appli
34	62	93.9	1516	15	US-10-060-036-166	Sequence 166, App
35	58	87.9	184	12	US-10-292-418-35	Sequence 35, Appl
36	58	87.9	184	15	US-10-131-241-49	Sequence 49, Appl
37	48	72.7	184	12	US-09-938-391-4	Sequence 4, Appli
38	48	72.7	230	12	US-09-938-391-2	Sequence 2, Appli
39	43	65.2	1867	11	US-09-824-574-5	Sequence 5, Appli
40	40	60.6	413	12	US-10-288-930-87	Sequence 87, Appl
41	40	60.6	556	15	US-10-128-714-3458	Sequence 3458, Ap
42	40	60.6	556	15	US-10-128-714-3584	Sequence 3584, Ap
43	40	60.6	576	15	US-10-128-714-8458	Sequence 8458, Ap
44	40	60.6	576	15	US-10-128-714-8584	Sequence 8584, Ap
45	39	59.1	181	15	US-10-149-819-14	Sequence 14, Appl

```
; TITLE OF INVENTION: Angiostatin and Endostatin Binding Proteins and Methods of Use
; FILE REFERENCE: 05213-0378 (43170-259333)
; CURRENT APPLICATION NUMBER: US/09/873,676
; CURRENT FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,065
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: US 60/289,387
; PRIOR FILING DATE: 2001-05-08
; NUMBER OF SEQ ID NOS: 123
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 115
; LENGTH: 20
; TYPE: PRT
; ORGANISM: mammalian
US-09-873-676-115

Query Match 100.0%; Score 66; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 3
US-09-405-499-1
; Sequence 1, Application US/09405499
; Patent No. US20020123458A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Endostatin Protein and Fragments Thereof
; FILE REFERENCE: 05213-0640
; CURRENT APPLICATION NUMBER: US/09/405,499
; CURRENT FILING DATE: 1999-09-23
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-405-499-1

Query Match 100.0%; Score 66; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 4
US-09-174-516-1
; Sequence 1, Application US/09174516A
; Patent No. US20020127595A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Methods of Detecting Endostatin Protein
; FILE REFERENCE: 05213-0227
; CURRENT APPLICATION NUMBER: US/09/174,516A
; CURRENT FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-09-174-516-1

Query Match 100.0%; Score 66; DB 10; Length 20;
```

```
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 5
US-09-154-302-1
; Sequence 1, Application US/09154302
; Patent No. US20020155987A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions
; TITLE OF INVENTION: and Methods
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/154,302
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/740,168
; FILING DATE: 22-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0223
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; TISSUE TYPE: Collagen
US-09-154-302-1

Query Match 100.0%; Score 66; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 6
US-10-351-284-1
; Sequence 1, Application US/10351284
; Publication No. US20030219426A1
; GENERAL INFORMATION:
```

```

; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions and Methods
; FILE REFERENCE: 05213-3051 (43170-282623)
; CURRENT APPLICATION NUMBER: US/10/351,284
; CURRENT FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/005,835
; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-02
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae gen. sp.
US-10-351-284-1

Query Match 100.0%; Score 66; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVHLHVALNTPL 13
Db 7 QPVHLHVALNTPL 19

RESULT 7
US-10-131-241-43
; Sequence 43, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Proliferation
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; CURRENT FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586
; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 43
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae sp.
US-10-131-241-43

Query Match 100.0%; Score 66; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVHLHVALNTPL 13
Db 7 QPVHLHVALNTPL 19

RESULT 8
US-10-232-316-1
; Sequence 1, Application US/10232316
; Publication No. US20030087393A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Therapeutic Antiangiogenic Compositions and Methods

```

```

; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jones & Askew, LLP
; STREET: 191 Peachtree, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/10/232,316
; FILING DATE: 03-Sep-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/740,168A
; FILING DATE: 22-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren, William L.
; REGISTRATION NUMBER: 36,714
; REFERENCE/DOCKET NUMBER: 05213-0223
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-818-3700
; TELEFAX: 404-818-3799
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Murine
; TISSUE TYPE: Collagen
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-232-316-1

Query Match 100.0%; Score 66; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVHLHVALNTPL 13
Db 7 QPVHLHVALNTPL 19

RESULT 9
US-10-042-347-1
; Sequence 1, Application US/10042347
; Publication No. US20030114370A1
; GENERAL INFORMATION:
; APPLICANT: O'Reilly, Michael S.
; APPLICANT: Folkman, M. Judah
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Endostatin Protein and Peptide Fragments
; FILE REFERENCE: 05213-0880 (43170-249874)
; CURRENT APPLICATION NUMBER: US/10/042,347
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: US 09/315,689
; PRIOR FILING DATE: 1999-05-20
; PRIOR APPLICATION NUMBER: US 60/106,343
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: US 09/154,302
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: US 08/740,168
; PRIOR FILING DATE: 1996-10-22
; PRIOR APPLICATION NUMBER: US 60/005,835

```


; PRIOR FILING DATE: 1995-10-23
; PRIOR APPLICATION NUMBER: US 60/023,070
; PRIOR FILING DATE: 1996-08-32
; PRIOR APPLICATION NUMBER: US 60/026,263
; PRIOR FILING DATE: 1996-09-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Murinae sp.
US-10-042-347-1

Query Match 100.0%; Score 66; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 7.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 10
US-10-292-418-18
; Sequence 18, Application US/10292418
; Publication No. US20030139365A1
; GENERAL INFORMATION:
; APPLICANT: Li, Yue
; APPLICANT: Gillies, Stephen D
; TITLE OF INVENTION: Expression and Export of Angiogenesis Inhibitors as
; FILE REFERENCE: LEX-006C1
; CURRENT APPLICATION NUMBER: US/10/292,418
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/383,315
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: US 60/097,893
; PRIOR FILING DATE: 1998-08-25
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 18
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-292-418-18

Query Match 100.0%; Score 66; DB 12; Length 184;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 11
US-10-131-241-46
; Sequence 46, Application US/10131241
; Publication No. US20030012792A1
; GENERAL INFORMATION:
; APPLICANT: Holaday, John W.
; APPLICANT: Fortier, Anne H.
; TITLE OF INVENTION: Compositions and Methods for Inhibiting Endothelial Cell Prolifer
; TITLE OF INVENTION: and Regulating Angiogenesis Using Cancer Markers
; FILE REFERENCE: 05213-0344 43170-271565
; CURRENT APPLICATION NUMBER: US/10/131,241
; PRIOR FILING DATE: 2002-07-22
; PRIOR APPLICATION NUMBER: US 09/413,049
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 09/316,802
; PRIOR FILING DATE: 1999-05-21
; PRIOR APPLICATION NUMBER: US 60/086,586

; PRIOR FILING DATE: 1998-05-22
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 46
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Murinae sp.
US-10-131-241-46

Query Match 100.0%; Score 66; DB 15; Length 184;
Best Local Similarity 100.0%; Pred. No. 0.00078;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTPL 13
Db 7 QPVLHLVALNTPL 19

RESULT 12
US-10-036-869-36
; Sequence 36, Application US/10036869
; Publication No. US20020151516A1
; GENERAL INFORMATION:
; APPLICANT: Mixson, James A
; TITLE OF INVENTION: CARRIER: DNA COMPLEXES CONTAINING DNA
; ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE
; THERAPY
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESS: Connolly, Bove, Lodge, & Hutz
; STREET: 1220 Market Street, P.O. Box 2207
; CITY: Wilmington
; STATE: Delaware
; COUNTRY: U.S.A.
; ZIP: 19899

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/036,869
FILING DATE: 29-NO. US20020151516A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/985,526
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/608,845
FILING DATE: 16-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: McMorro Jr., Robert G
TELEPHONE: (302) 658-9141
TELEFAX: (302) 658-5613
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 185 amino acids
TYPE: amino acid
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 36:
US-10-036-869-36

Query Match 100.0%; Score 66; DB 14; Length 185;
Best Local Similarity 100.0%; Pred. No. 0.00079;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTPL 13
Db 8 QPVLHLVALNTPL 20

RESULT 13
US-09-998-831-13

Sequence 13, Application US/09998831
 Patent No. US2002011915A1
 GENERAL INFORMATION:
 APPLICANT: Philip E. Thorpe
 APPLICANT: Rolf A. Brekken
 TITLE OF INVENTION: ANTIBODY CONJUGATE COMPOSITIONS FOR SELECTIVELY
 TITLE OF INVENTION: INHIBITING VEGF
 FILE REFERENCE: 4001.002584
 CURRENT APPLICATION NUMBER: US/09/998,831
 CURRENT FILING DATE: 2003-11-30
 PRIOR APPLICATION NUMBER: 09/561,108
 PRIOR FILING DATE: 2000-04-28
 NUMBER OF SEQ ID NOS: 44
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 13
 LENGTH: 191
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
 US-09-998-831-13

Query Match 100.0%; Score 66; DB 10; Length 191;
 Best Local Similarity 100.0%; Pred. No. 0.00082;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTPL 13
 |||||
 Db 14 QPVLHLVALNTPL 26

RESULT 14

US-10-373-561-13
 Sequence 13, Application US/10373561
 Publication No. US20030175276A1
 GENERAL INFORMATION:
 APPLICANT: Philip E. Thorpe
 APPLICANT: Rolf A. Brekken
 TITLE OF INVENTION: ANTIBODY METHODS FOR SELECTIVELY INHIBITING VEGF
 FILE REFERENCE: 4001.002582
 CURRENT APPLICATION NUMBER: US/10/373,561
 CURRENT FILING DATE: 2003-02-24
 PRIOR APPLICATION NUMBER: US/09/561,499
 PRIOR FILING DATE: 2000-04-28
 PRIOR APPLICATION NUMBER: 60/131,432
 PRIOR FILING DATE: 1999-04-28
 NUMBER OF SEQ ID NOS: 44
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 13
 LENGTH: 191
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: SYNTHETIC
 US-10-373-561-13

Query Match 100.0%; Score 66; DB 12; Length 191;
 Best Local Similarity 100.0%; Pred. No. 0.00082;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTPL 13
 |||||
 Db 14 QPVLHLVALNTPL 26

RESULT 15

US-10-422-934-71
 Sequence 71, Application US/10422934
 Publication No. US20030186841A1
 GENERAL INFORMATION:
 APPLICANT: Barbas, Carlos F., III
 APPLICANT: Kadan, Michael
 APPLICANT: Beerli, Roger

TITLE OF INVENTION: LIGAND ACTIVATED TRANSCRIPTIONAL REGULATOR PROTEINS
 FILE REFERENCE: 22908-1227C
 CURRENT APPLICATION NUMBER: US/10/422,934
 CURRENT FILING DATE: 2003-04-23
 PRIOR APPLICATION NUMBER: 09/586,625
 PRIOR FILING DATE: 2000-06-02
 PRIOR APPLICATION NUMBER: 09/433,042
 PRIOR FILING DATE: 1999-10-25
 NUMBER OF SEQ ID NOS: 92
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 71
 LENGTH: 207
 TYPE: PRT
 ORGANISM: Muridae
 US-10-422-934-71

Query Match 100.0%; Score 66; DB 12; Length 207;
 Best Local Similarity 100.0%; Pred. No. 0.00089;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QPVLHLVALNTPL 13
 |||||
 Db 30 QPVLHLVALNTPL 42

Search completed: January 26, 2004, 19:16:33
 Job time : 368 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2004, 19:04:50 ; Search time 38 Seconds
(without alignments)
32.900 Million cell updates/sec

Title: US-09-766-412-30

Perfect score: 66

Sequence: 1 QPVLHLVALNPL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	100.0	1315	2 A56101	collagen alpha 1(XVIII)
2	66	100.0	1774	2 B56101	collagen alpha 1(X)
3	62	93.9	684	2 A3019	collagen alpha 1(X)
4	44	66.7	1388	2 A3317	collagen alpha 1(X)
5	43	65.2	1867	2 S2275	MOT1 protein - yea
6	40	60.6	176	2 C70602	hypothetical prote
7	40	60.6	342	2 A1045	conserved hypothet
8	39	59.1	190	2 D75319	3-methyladenine gl
9	39	59.1	333	2 I40791	acetoin dehydrogen
10	39	59.1	425	2 A80796	NADH2 dehydrogenas
11	39	59.1	533	2 G96541	probable cytochrom
12	38	57.6	103	2 T50934	hypothetical prote
13	38	57.6	160	2 E72631	hypothetical prote
14	38	57.6	258	2 T25958	hypothetical prote
15	38	57.6	336	2 T4958	hypothetical prote
16	38	57.6	617	2 T15104	hypothetical prote
17	38	57.6	2311	2 T06161	acetyl-CoA carboxy
18	38	57.6	3206	1 GNVSVU	genome polyprotein
19	37	56.1	102	2 S0397	probable membrane
20	37	56.1	191	1 QBE4	BHRF1 protein - hu
21	37	56.1	519	2 A36542	probable cytochrom
22	37	56.1	728	1 TVHUSK	transforming prote
23	36	54.5	72	1 PNBPF6	p13 protein - phag
24	36	54.5	298	2 T22264	hypothetical prote
25	36	54.5	323	2 T48425	lipase-like protei
26	36	54.5	331	2 A83534	probable C4-dicarb
27	36	54.5	342	1 S56374	hypothetical 38.7K
28	36	54.5	342	2 G31269	hypothetical prote
29	36	54.5	342	2 E86110	hypothetical prote

30 36 54.5 358 2 T15069
31 36 54.5 370 2 H72748
32 36 54.5 548 2 G82286
33 36 54.5 651 2 AG0977
34 36 54.5 656 2 C91186
35 36 54.5 656 2 B86033
36 36 54.5 1162 2 T37889
37 35 53.0 111 2 F83574
38 35 53.0 190 1 PWECK
39 35 53.0 190 2 B85569
40 35 53.0 190 2 D90719
41 35 53.0 193 2 A81149
42 35 53.0 133 2 H81874
43 35 53.0 195 2 A80587
44 35 53.0 222 2 H71260
45 35 53.0 222 2 B72729

ALIGNMENTS

RESULT 1

A56101
collagen alpha 1(XVIII) chain precursor, short splice form - mouse
N;Contains: endostatin
C;Species: Mus musculus (house mouse)
C;Date: 03-Oct-1995 #sequence_revision 08-May-1998 #text_change 31-Mar-2000
C;Accession: A56101; A58371; S72450; S65595
R;Rehn, M.; Pihlajaniemi, T.
J. Biol. Chem. 270, 4705-4711, 1995
A;Title: Identification of three N-terminal ends of type XVIII collagen chains and tissue
tif homologous to rat and Drosophila frizzled proteins.
A;Reference number: A56101; MUID:95181468; PMID:7876242
A;Accession: A56101
A;Molecule type: mRNA
A;Residues: 1-103 <REH1>
A;Cross-references: GB:U11636; NID:9618427; PIDN:AAC52178.1; PID:G618428
R;Rehn, M.; Pihlajaniemi, T.
Proc. Natl. Acad. Sci. U.S.A. 91, 4234-4238, 1994
A;Title: Alpha1(XVIII), a collagen chain with frequent interruptions in the collagenous
A;Reference number: A58371; MUID:94240112; PMID:8183894
A;Accession: A58371
A;Molecule type: mRNA
A;Residues: 1-928 <REH2>
A;Cross-references: GB:I116898; NID:9404754; PIDN:AAA37434.1; PID:G553894
R;Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.
submitted to the EMBL Data Library, August 1993
A;Reference number: S72450
A;Accession: S72450
A;Molecule type: mRNA
A;Residues: 28-687, 'L', '689-734', 'F', '736-751', 'R', '753-1315' <ORH>
A;Cross-references: EMBL:L22545; NID:G348968; PIDN:AAA19787.1; PID:G511298
R;Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, A.B.R.
Proc. Natl. Acad. Sci. U.S.A. 91, 4229-4233, 1994
A;Title: Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-Xaa-
A;Reference number: A58370; MUID:94240111; PMID:8183893
A;Accession: S65595
A;Molecule type: mRNA
A;Residues: 28-1315 <OHS>
A;Cross-references: EMBL:L22545
C;Comment: Prolines and lysines at the third position of the tripeptide repeating unit (C
lated and subsequently O-glycosylated.
C;Comment: The different splice forms of collagen alpha 1(XVIII) may be involved in peri
C;Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of un
ay be useful in treating solid tumors.
C;Genetics:
A;Gene: MGI:Col18a1
A;Cross-references: MGI:71175
A;Map position: 10:41.0
C;Superfamily: unassigned collagens
C;Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteoglyc
F;1-25/Domain: signal sequence #status predicted <SIG>
F;24-235/Region: thrombospondin amino-terminal similarity

F:26-1315/Product: collagen alpha 1(XVIII) chain, short splice form #status predicted <M>
 F:327-353/Domain: collagenous #status predicted <CO1>
 F:364-437/Domain: collagenous #status predicted <CO2>
 F:462-583/Domain: collagenous #status predicted <CO3>
 F:607-689/Domain: collagenous #status predicted <CO4>
 F:704-745/Domain: collagenous #status predicted <CO5>
 F:759-831/Domain: collagenous #status predicted <CO6>
 F:842-874/Domain: collagenous #status predicted <CO7>
 F:887-910/Domain: collagenous #status predicted <CO8>
 F:892-894/Region: cell attachment (R-G-D) motif
 F:918-969/Domain: collagenous #status predicted <CO9>
 F:983-1000/Domain: collagenous #status predicted <CO10>
 F:1132-1315/Product: endostatin #status predicted <EST>
 F:1139-1315/Region: multiplexin collagen carboxyl-terminal similarity
 F:126-498/Binding site: carboxydrate (Asn) (covalent) #status predicted
 F:172-228/Disulfide bonds: #status predicted
 F:240-454,1257/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted
 F:451,454,594/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 100.0%; Score 66; DB 2; Length 1315;
 Best Local Similarity 100.0%; Pred. No. 0.00055; Indels 0; Gaps 0;
 Matches 13; Conservative 0; Mismatches 0;

QY 1 QPVLHLVALNTPPL 13
 |||||
 DB 1138 QPVLHLVALNTPPL 1150

RESULT 2
 B56101
 N:Contains: collagen alpha 1(XVIII) chain precursor, long splice form - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Oct-1995 #sequence_revision 08-May-1998 #text_change 31-Mar-2000
 C:Accession: B56101; C56101; S72450; S65595; PNO675; A54072; A58816
 R:Rehn, M.; Pihlajaniemi, T.
 J. Biol. Chem. 270, 4705-4711, 1995
 A:Title: Identification of three N-terminal ends of type XVIII collagen chains and tissue homologous to rat and Drosophila frizzled proteins.
 A:Reference number: A56101; MUID:95181468; PMID:7876242
 A:Accession: B56101
 A:Molecule type: mRNA
 A:Residues: 1-362 <REH1>
 A:Cross-references: GB:U11637; NID:G618430; PIDN:AA52179.1; PID:G618430
 A:Experimental source: splice form clone PE17.24
 A:Accession: C56101
 A:Molecule type: mRNA
 A:Residues: 1-239,487-562 <REH2>
 A:Cross-references: GB:U11637; NID:G618429
 A:Experimental source: splice form clones PB8.1, PB19, PB15.2
 R:Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, B.R.
 submitted to the EMBL Data Library, August 1993
 A:Reference number: S72450
 A:Accession: S72450
 A:Molecule type: mRNA
 A:Residues: 487-1146 'L', 1148-1193 'F', 1195-1210 'R', 1212-1512 'L', 1514-1522 'F', 1524-1616 'R'
 A:Cross-references: EMBL:L22545; NID:G348968; PIDN:AA19787.1; PID:G511298
 R:Oh, S.P.; Kamagata, Y.; Muragaki, Y.; Timmons, S.; Ooshima, A.; Olsen, A.B.R.
 Proc. Natl. Acad. Sci. U.S.A. 91, 4229-4233, 1994
 A:Title: Isolation and sequencing of cDNAs for proteins with multiple domains of Gly-Xaa
 A:Reference number: A58370; MUID:94240111; PMID:8183893
 A:Accession: S65595
 A:Molecule type: mRNA
 A:Residues: 487-1512 'L', 1514-1522 'F', 1524-1693 'V', 1695-1774 <OH2>
 A:Cross-references: EMBL:L22545
 R:Abé, N.; Muragaki, Y.; Yoshioka, H.; Inoue, H.; Ninomiya, Y.
 Biochem. Biophys. Res. Commun. 196, 576-582, 1993
 A:Title: Identification of a novel collagen chain represented by extensive interruptions
 A:Reference number: PNO675; MUID:94059075; PMID:8240330
 A:Accession: PNO675
 A:Molecule type: mRNA
 A:Residues: 635-1774 <AB>
 R:Rehn, M.; Hintikka, E.; Pihlajaniemi, T.

J. Biol. Chem. 269, 13929-13935, 1994
 A:Title: Primary structure of the alpha chain of mouse type XVIII collagen, partial str collagen chain.
 A:Reference number: A54072; MUID:94245707; PMID:8188673
 A:Accession: A54072
 A:Molecule type: DNA; mRNA
 A:Residues: 1293-1403, R', 1405-1774 <REH3>
 A:Cross-references: GB:U03714; NID:9487733; PIDN:AAA20657.1; PID:9487734
 R:O'Reilly, M.S.; Boehm, T.; Shing, Y.; Fukai, N.; Vasios, G.; Lane, W.S.; Flynn, E.; Bjiell, 88, 277-285, 1997
 A:Title: Endostatin: an endogenous inhibitor of angiogenesis and tumor growth.
 A:Reference number: A58816; MUID:97160848; PMID:9008168
 A:Accession: A58816
 A:Molecule type: Protein
 A:Residues: 1591-1610 <OR>
 A:Experimental source: hemangioendothelium cells
 A:Note: Inhibits endothelial cell proliferation
 C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit (labeled and subsequently O-glycosylated).
 C:Comment: The different splice forms of collagen alpha 1(XVIII) may be involved in pericellular matrix remodeling.
 C:Comment: Endostatin is released from collagen alpha 1(XVIII) chain by the action of uric acid.
 C:Genetics:
 A:Gene: MGI:Coll18a1
 A:Cross-references: MGI:71175
 A:Map position: 10:41.0
 A:Introns: 1295/3; 1310/1; 1331/1; 1345/3; 1388/3; 1437/1; 1461/3; 1505/3; 1516/3; 1599/3
 A:Note: the list of introns is incomplete
 C:Superfamily: unassigned collagens
 C:Keywords: alternative splicing; angiogenesis inhibitor; chondroitin sulfate proteoglycan precursor; collagen alpha 1(XVIII) chain precursor, long splice form #status predicted <CO1>
 F:1-239,487-1774/Product: collagen alpha 1(XVIII) chain precursor, medium splice form #status predicted <CO2>
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:361-486/Region: frizzled similarity
 F:786-812/Domain: collagenous #status predicted <CO01>
 F:823-896/Domain: collagenous #status predicted <CO02>
 F:921-1042/Domain: collagenous #status predicted <CO03>
 F:1066-1148/Domain: collagenous #status predicted <CO04>
 F:1163-1204/Domain: collagenous #status predicted <CO05>
 F:1218-1290/Domain: collagenous #status predicted <CO06>
 F:1301-1333/Domain: collagenous #status predicted <CO07>
 F:1346-1369/Domain: collagenous #status predicted <CO08>
 F:1351-1353/Region: cell attachment (R-G-D) motif
 F:1377-1428/Domain: collagenous #status predicted <CO09>
 F:1442-1459/Domain: collagenous #status predicted <CO10>
 F:1591-1774/Product: endostatin #status predicted <EST>
 F:1598-1774/Region: multiplexin collagen carboxyl-terminal similarity
 F:354,361,947/Binding site: carboxydrate (Asn) (covalent) #status predicted
 F:599,704,1716/Binding site: carboxydrate (Ser) (covalent) #status predicted
 F:910,913,1053/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 100.0%; Score 66; DB 2; Length 1774;
 Best Local Similarity 100.0%; Pred. No. 0.00076; Indels 0; Gaps 0;
 Matches 13; Conservative 0; Mismatches 0;

QY 1 QPVLHLVALNTPPL 13
 |||||
 DB 1597 QPVLHLVALNTPPL 1609

RESULT 3
 A53019
 collagen alpha 1(XVIII) chain - human (fragment)
 N:Contains: endostatin
 C:Species: Homo sapiens (man)
 C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 31-Mar-2000
 C:Accession: A53019
 R:Oh, S.P.; Warman, M.L.; Seidlin, M.F.; Cheng, S.D.; Knoll, J.H.M.; Timmons, S.; Olsen, G.
 Genomics 19, 494-499, 1994
 A:Title: Cloning of cDNA and genomic DNA encoding human type XVIII collagen and localiza
 A:Reference number: A53019; MUID:94245237; PMID:8188291
 A:Accession: A53019
 A:Molecule type: mRNA

A;Map position: 9q21-9q22
C;Superfamily: unassigned collagens
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-1388/Product: collagen alpha 1(XV) chain #status predicted <MAT>
F;1216-1388/Region: multiplexin collagen carboxyl-terminal similarity

Query Match 66.7%; Score 44; DB 2; Length 1388;
Best Local Similarity 66.7%; Pred. No. 8.3;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTP 12
: ||| |||
Db 1215 KPALHLAALNMP 1226

RESULT 5
S22775
MOT1 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein LPF4c; protein YPL082c
C;Species: Saccharomyces cerevisiae
C;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C;Accession: S22775; S61106
R;Davis, J.L.; Kunisawa, R.; Thorner, J.
Mol. Cell. Biol. 12, 1879-1892, 1992
A;Title: A presumptive helicase (MOT1 gene product) affects gene expression and is required for normal growth of *Saccharomyces cerevisiae*
A;Reference number: S22775; MUID:92195335; PMID:1312673
A;Accession: S22775
A;Molecule type: DNA
A;Residues: 1-1867 <DAV>
A;Cross-references: EMBL:M83224; NID:G171964; PIDN:AAA34786.1; PID:G171965
R;Hall, J.; Ahmed, A.; Bussey, H.; Fortin, N.; Priesen, J.D.; Storms, R.K.; War
submitted to the EMBL Data Library, August 1995
A;Description: The sequence of *Saccharomyces cerevisiae* chromosome XVI left arm.
A;Reference number: S59677
A;Accession: S61106
A;Molecule type: DNA
A;Residues: 1-1867 <HAL>
A;Cross-references: EMBL:U41849; NID:G1147608; PID:G1147612; MIPS:YPL082c
C;Genetics:
A;Gene: SGD:MOT1
A;Cross-references: SGD:S0006003; MIPS:YPL082c
A;Map position: 16L
C;Keywords: DNA binding; nucleus; transmembrane protein
F;700-716/Domain: transmembrane #status predicted <TM1>
F;1038-1054/Domain: transmembrane #status predicted <TM2>
F;1186-1202/Domain: transmembrane #status predicted <TM3>

Query Match 65.2%; Score 43; DB 2; Length 1867;
Best Local Similarity 69.2%; Pred. No. 18;
Matches 9; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

QY 1 QPVLHLVALNTP 13
: ||| |||
Db 637 QPILHL--LNTPV 647

RESULT 6
C70602
hypothetical protein Rv1000 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C;Accession: C70602
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Scares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A;Reference number: A70500; MUID:96295987; PMID:9634230
A;Accession: C70602
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-176 <COL>

A:Cross-references: GB:294752; GB:AL123456; NID:G3261731; PIDN:CAB08143.1; PID:G2052135
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: Rv1000
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv1000

Query Match 60.6%; Score 40; DB 2; Length 176;
Best Local Similarity 63.6%; Pred. No. 4.8;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PVLHLVALNTP 12
|||:|||||
DB 103 PWHLTALGSP 113
|||:|||||

RESULT 7
A:Accession: AH1045
A:Title: conserved hypothetical protein yJek [imported] - Salmonella enterica subsp. enterica serovar Typhi
A:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AH1045
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AH1045
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-342 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD06813.1; PID:G16505463; GSPDB:GN00176
C:Genetics:
A:Gene: yJek
C:Superfamily: conserved hypothetical protein yodo

Query Match 60.6%; Score 40; DB 2; Length 342;
Best Local Similarity 77.8%; Pred. No. 10;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 HVALNTP 13
|||:|||||
DB 3 HIVTUNTP 11
|||:|||||

RESULT 8
D75319
3-methyladenine glycosidase - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 15-Jun-2001
C:Accession: D75319
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: D75319
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-190 <WHI>
A:Cross-references: GB:AE002043; GB:AE000513; NID:G6459859; PIDN:AAFL1623.1; PID:G645986
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2074
C:Superfamily: Bacillus subtilis DNA-3-methyladenine glycosidase homolog yxjJ

Query Match 59.1%; Score 39; DB 2; Length 190;
Best Local Similarity 66.7%; Pred. No. 8.1;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PVLHLVALNTP 13
|||:|||||
DB 136 PELHLLAPETPL 147
|||:|||||

RESULT 9
I40791
acetoin dehydrogenase (TPP-dependent) (EC 1.-.-.-) beta chain - Clostridium magnum
C:Species: Clostridium magnum
C:Date: 04-Sep-1997 #sequence_revision 04-Sep-1997 #text_change 11-Jun-1999
C:Accession: I40791
R:Kruger, N.; Oppermann, F.B.; Lorenzl, H.; Steinbuechel, A.
J. Bacteriol. 176, 3614-3630, 1994
A:Title: Biochemical and molecular characterization of the Clostridium magnum acetoin dehydrogenase
A:Reference number: I40789; MUID:94266715; PMID:8206840
A:Accession: I40791
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-333 <XRU>
A:Cross-references: GB:I31844; NID:G473324; PIDN:AAA21745.1; PID:G47327
C:Superfamily: pyruvate dehydrogenase (lipoamide) beta chain
C:Keywords: oxidoreductase

Query Match 59.1%; Score 39; DB 2; Length 333;
Best Local Similarity 58.3%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 PVLHLVALNTP 13
|||:|||||
DB 297 PVRIGALNTP 308
|||:|||||

RESULT 10
AB0796
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) - Salmonella enterica subsp. enterica serovar Typhi
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AB0796
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AB0796
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-425 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD07548.1; PID:G16503540; GSPDB:GN00176
C:Genetics:
A:Gene: STY2546
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 2
C:Keywords: oxidoreductase

Query Match 59.1%; Score 39; DB 2; Length 425;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTP 13
|||:|||||
DB 412 QPLISLVQATPL 424
|||:|||||

RESULT 11
G96541
probable cytochrome P450 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 20-Apr-2001
C:Accession: G96541
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 815-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: G96541
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-533 <STO>
A:Cross-references: GB:AB005173; NID:g11054632; PIDN:AAG27877.1; GSPDB:GN00141
C:Genetics:
A:Gene: F1J76.6
A:Map position: 1
A:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: heme; iron; metalloprotein
F:457/Binding site: heme iron (Cys) (axial ligand) #status predicted
Query Match 59.1%; Score 39; DB 2; Length 533;
Best Local Similarity 58.3%; Pred. No. 25; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 2 PVLHLVALNTPL 13
Db 77 PLLHLAFNIPI 88
RESULT 12
T50954
hypothetical protein B24P7.90 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 08-Sep-2000
A:Accession: T50954
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T50954
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-103 <SCH>
A:Cross-references: EMBL:AL398890; GSPDB:GN00116; NCSP:B24P7.90
A:Experimental source: BAC clone B24P7; strain OR74A
C:Genetics:
A:Gene: NCSP:B24P7.90
A:Map position: 6
A:Superfamily: Neurospora crassa hypothetical protein B24P7.90
Query Match 57.6%; Score 38; DB 2; Length 103;
Best Local Similarity 58.3%; Pred. No. 6.4; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
Qy 1 QPVLHLVALNTPL 12
Db 70 QPVLHLQLESP 81
RESULT 13
E72631
hypothetical protein APE1508 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
A:Accession: E72631
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hakiawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: E72631
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-160 <KAW>
A:Cross-references: DDBJ:AP000061; NID:g5104821; PIDN:BAA80507.1; PID:d1044293; PID:g5101
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1508
Query Match 57.6%; Score 38; DB 2; Length 160;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
Qy 2 PVLHLVALNTPL 13
Db 56 PVLHLPLDIPI 67
RESULT 14
T25958
hypothetical protein ZC204.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
A:Accession: T25958
R:Wamsley, P.; Kramer, J.
submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid ZC204.
A:Reference number: Z20116
A:Accession: T25958
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-258 <WAM>
A:Cross-references: EMBL:U80839; PIDN:AA837912.1; GSPDB:GN00020; CESP:ZC204.3
A:Experimental source: strain Bristol N2; clone ZC204
C:Genetics:
A:Gene: CESP:ZC204.3
A:Map position: 2
A:Introns: 5/1; 123/3; 167/1
Query Match 57.6%; Score 38; DB 2; Length 258;
Best Local Similarity 80.0%; Pred. No. 18; Indels 0; Gaps 0;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 4 LHLVALNTPL 13
Db 85 LYLVALETPL 94
RESULT 15
T44958
hypothetical protein [imported] - Natronomonas pharaonis
C:Species: Natronomonas pharaonis
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
A:Accession: T44958
R:Mattar, S.; Souquet, M.; Heinrich, H.J.; Engelhard, M.
submitted to the EMBL Data Library, August 1996
A:Description: The first fully resolved primary structure of an archaeal succinate-dehydi
A:Reference number: Z22881
A:Accession: T44958
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-336 <MAT>
A:Cross-references: EMBL:Y07709; PIDN:CAA68978.1
A:Experimental source: strain SP1 /28
Query Match 57.6%; Score 38; DB 2; Length 336;
Best Local Similarity 63.6%; Pred. No. 24; Indels 0; Gaps 0;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 2 PVLHLVALNTPL 12
Db 312 PLCHLVALDEP 322
Search completed: January 26, 2004, 19:08:54

Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2004, 19:04:50 ; Search time 23 Seconds
(without alignments)
26,580 Million cell updates/sec

Title: US-09-766-412-30
Perfect score: 66
Sequence: 1 QPVLHVALNTPPL 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	100.0	1527	1 CAIH_MOUSE	P39061 mus musculus
2	62	93.9	1516	1 CAIH_HUMAN	P39060 homo sapien
3	44	66.7	1388	1 CAIE_HUMAN	P39059 homo sapien
4	43	65.2	1867	1 MOT1_YEAST	P32333 saccharomyc
5	40	60.6	217	1 ATP6_RHODP	Q36836 rhopalosiph
6	39	59.1	190	1 MGH_DEIRA	Q39630 deinococcus
7	38	57.6	3206	1 POLG_PSNV	P29152 p genome po
8	37	56.1	102	1 YN21_YEAST	P40211 saccharomyc
9	37	56.1	191	1 EAR_EBV	P01182 Epstein-bar
10	37	56.1	348	1 SKI_MOUSE	Q60598 mus musculus
11	37	56.1	728	1 SKI_HUMAN	P12755 homo sapien
12	36	54.5	72	1 VP13_BPH6	P11130 bacterioph
13	36	54.5	323	1 UCP4_HUMAN	Q95847 homo sapien
14	36	54.5	342	1 YUEK_ECOLI	P39280 escherichia
15	36	54.5	370	1 EGSA_AERPE	Q9yer2 aeropyrum p
16	36	54.5	749	1 NIBL_MOUSE	Q8r1f1 mus musculus
17	35	53.0	190	1 ATKC_ECO57	Q8x9G0 escherichia
18	35	53.0	190	1 ATKC_ECOL6	Q8f1v5 escherichia
19	35	53.0	190	1 ATKC_ECOLI	P03961 escherichia
20	35	53.0	193	1 LOLE_NEIMA	P57023 neisseria m
21	35	53.0	193	1 LOLE_NEIMB	P7024 reisseria m
22	35	53.0	194	1 ATKC_SALTI	Q8z866 salmonella
23	35	53.0	194	1 ATKC_SALTY	Q8z9G3 salmonella
24	35	53.0	222	1 GIDB_TREPA	Q66106 treponema p
25	35	53.0	329	1 ANFK_CLOHU	Q46084 clostridium
26	35	53.0	332	1 G3P_EHYIN	P26988 phytophthor
27	35	53.0	348	1 ARC2_DEIRA	Q3rvd3 deinococcus
28	35	53.0	387	1 CYB_SCHPO	P05501 schizosacch
29	35	53.0	397	1 LIPG_BOVIN	Q29458 bos taurus
30	35	53.0	462	1 ANFK_AZOVI	P16267 azotobacter
31	35	53.0	618	1 ORC2_DROME	Q24168 drosophila
32	35	53.0	623	1 HCVE_EURCA	P02242 eurytelma c
33	35	53.0	956	1 RRPO_EBNV	P21405 southern be

RESULT 1

ID	CAIH_MOUSE	STANDARD;	PRT;	1527 AA.
AC	P39061; Q61437; Q62002;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-SEP-2003 (Rel. 42, Last annotation update)			
DE	Collagen alpha 1(XVIII) chain precursor (Contains: Endostatin).			
GN	COL18A1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM SHORT).			
RC	STRAIN=BALB/c; TISSUE=Liver;			
RX	MEDLINE=94245707; PubMed=8188673;			
RA	Rehn M.V., Hintikka E., Pihlajaniemi T.;			
RT	"Primary structure of the alpha 1 chain of mouse type XVIII collagen,			
RT	partial structure of the corresponding gene, and comparison of the			
RT	alpha 1(XVIII) chain with its homologue, the alpha 1(XV) collagen			
RT	chain.";			
RL	J. Biol. Chem. 269:13929-13935(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).			
RA	Rehn M., Hintikka E., Pihlajaniemi T.;			
RT	"Characterization of the mouse gene for the alpha-1 chain of type			
RT	XVIII collagen (COL18A1) reveals that the three variant N-terminal			
RT	polypeptide forms are transcribed from two widely separated			
RT	promoters.";			
RT	Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.			
RL	Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE OF 213-1140 FROM N.A. (ISOFORM SHORT).			
RX	MEDLINE=94240112; PubMed=8183894;			
RA	Rehn M.V., Pihlajaniemi T.;			
RT	"Alpha 1(XVIII), a collagen chain with frequent interruptions in the			
RT	collagenous sequence, a distinct tissue distribution, and homology			
RT	with type XV collagen.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 91:4234-4238(1994).			
RN	[4]			
RP	SEQUENCE OF 240-1527 FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=94240111; PubMed=8183893;			
RA	Oh S.P., Kamagata Y., Muragaki Y., Timmons S., Coshima A., Olsen B.R.;			
RT	"Isolation and sequencing of cDNAs for proteins with multiple domains			
RT	of Gly-Xaa-Yaa repeats identify a distinct family of collagenous			
RT	proteins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 91:4229-4233(1994).			
RN	[5]			
RP	CHARACTERIZATION OF ENDOSTATIN, AND PARTIAL SEQUENCE.			
RA	MEDLINE=97160848; PubMed=9008168;			
RX	O'Reilly M.S., Boehm T., Shing Y., Fukai N., Vasios G., Lane W.S.,			
RA	Flynn E., Birkhead J.R., Olsen B.R., Folkman J.;			
RT	"Endostatin: an endogenous inhibitor of angiogenesis and tumor			
RT	growth.";			
RL	Cell 88:277-285(1997).			

P12689 saccharomyc
O83273 treponema p
P02948 rhodobacter
P02949 rhodobacter
P07116 escherichia
Q8u9e0 agrobacteri
Q9v1y5 pyrococcus
Q00275 apis mellif
Q9b1y9 homo sapien
P30138 escherichia
P33832 variola vir
P07193 xenopus lae

ALIGNMENTS

[6]
X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF ENDOSTATIN.
MEDLINE=98169387; PubMed=9501087;
Hohenester E., Sasaki T., Olsen B.R., Timpl R.;
"Crystal structure of the angiogenesis inhibitor endostatin at 1.5-A
resolution.";
EMBO J. 17:1656-1664(1998).
CC -!- FUNCTION: ENDOSTATIN POTENTLY INHIBITS ENDOTHELIAL CELL
CC PROLIFERATION AND ANGIOGENESIS. MAY INHIBIT ANGIOGENESIS BY
CC BINDING TO THE HEPARAN SULPHATE PROTEOGLYCANS INVOLVED IN GROWTH
CC FACTOR SIGNALLING.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Long;
CC IsoId=P39061-1; Sequence=Displayed;
CC Name=Short;
CC IsoId=P39061-2; Sequence=VSP_001157, VSP_001158;
CC -!- PMW: Prolines at the third position of the tripeptide repeating
CC unit (G-X-Y) are hydroxylated in some or all of the chains.
CC -!- SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH
CC INTERRUPTED HELICES (FACIT) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; L16898; AAC37434.1; -;
EMBL; U03714; AAC20657.1; -;
EMBL; U03715; AAC52901.1; -;
EMBL; U34606; AAC52901.1; JOINED.
EMBL; U34608; AAC52901.1; JOINED.
EMBL; U34609; AAC52901.1; JOINED.
EMBL; U34610; AAC52901.1; JOINED.
EMBL; U34611; AAC52901.1; JOINED.
EMBL; U34612; AAC52901.1; JOINED.
EMBL; U34613; AAC52901.1; JOINED.
EMBL; U03716; AAC52902.1; JOINED.
EMBL; U03717; AAC52901.1; JOINED.
EMBL; U03718; AAC52901.1; JOINED.
EMBL; U03715; AAC52902.1; -;
EMBL; U34607; AAC52902.1; JOINED.
EMBL; U34608; AAC52902.1; JOINED.
EMBL; U34609; AAC52902.1; JOINED.
EMBL; U34610; AAC52902.1; JOINED.
EMBL; U34611; AAC52902.1; JOINED.
EMBL; U34612; AAC52902.1; JOINED.
EMBL; U34613; AAC52902.1; JOINED.
EMBL; U03716; AAC52902.1; JOINED.
EMBL; U03718; AAC52902.1; JOINED.
EMBL; U11636; AAC52178.1; -;
EMBL; L22545; AAA19787.1; -;
PIR; A56101; A56101.
PDB; 1KOE; 16-FEB-99.
PDB; 1DYO; 11-APR-00.
PDB; 1DYI; 21-JAN-01.
MGD; MGI:88451; Coll18a1.
GO; GO:0001525; P:angiogenesis; IMP.
InterPro; IPR000087; Collagen.
InterPro; IPR001791; Laminin_G.
InterPro; IPR003129; TSPN.
Pfam; PF01391; Collagen; 8.
Pfam; PF02210; TSPN; 1.
ProDom; PDO00007; C1g_helix; 1.
SMART; SM00282; LamG; 1.
SMART; SM00210; TSPN; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing;
3D-structure.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 1527 COLLAGEN ALPHA 1(XVIII) CHAIN.

1344 1527
CHAIN 27
DOMAIN 538
DOMAIN 539
DOMAIN 566
DOMAIN 575
DOMAIN 649
DOMAIN 673
DOMAIN 795
DOMAIN 818
DOMAIN 901
DOMAIN 915
DOMAIN 957
DOMAIN 958
DOMAIN 970
DOMAIN 1043
DOMAIN 1053
DOMAIN 1054
DOMAIN 1086
DOMAIN 1087
DOMAIN 1089
DOMAIN 1122
DOMAIN 1123
DOMAIN 1129
DOMAIN 1130
DOMAIN 1181
DOMAIN 1182
DOMAIN 1194
DOMAIN 1195
DOMAIN 1212
DOMAIN 1213
DOMAIN 1527
CARBOHYD 338
CARBOHYD 338
DISULFID 700
DISULFID 1376
SITE 1478
SITE 1508
SITE 1516
SITE 1518
SITE 1519
SITE 1520
SITE 1521
SITE 1522
SITE 1523
SITE 1524
SITE 1525
SITE 1526
SITE 1527
SITE 1528
SITE 1529
SITE 1530
SITE 1531
SITE 1532
SITE 1533
SITE 1534
SITE 1535
SITE 1536
SITE 1537
SITE 1538
SITE 1539
SITE 1540
SITE 1541
SITE 1542
SITE 1543
SITE 1544
SITE 1545
SITE 1546
SITE 1547
SITE 1548
SITE 1549
SITE 1550
SITE 1551
SITE 1552
SITE 1553
SITE 1554
SITE 1555
SITE 1556
SITE 1557
SITE 1558
SITE 1559
SITE 1560
SITE 1561
SITE 1562
SITE 1563
SITE 1564
SITE 1565
SITE 1566
SITE 1567
SITE 1568
SITE 1569
SITE 1570
SITE 1571
SITE 1572
SITE 1573
SITE 1574
SITE 1575
SITE 1576
SITE 1577
SITE 1578
SITE 1579
SITE 1580
SITE 1581
SITE 1582
SITE 1583
SITE 1584
SITE 1585
SITE 1586
SITE 1587
SITE 1588
SITE 1589
SITE 1590
SITE 1591
SITE 1592
SITE 1593
SITE 1594
SITE 1595
SITE 1596
SITE 1597
SITE 1598
SITE 1599
SITE 1600
SITE 1601
SITE 1602
SITE 1603
SITE 1604
SITE 1605
SITE 1606
SITE 1607
SITE 1608
SITE 1609
SITE 1610
SITE 1611
SITE 1612
SITE 1613
SITE 1614
SITE 1615
SITE 1616
SITE 1617
SITE 1618
SITE 1619
SITE 1620
SITE 1621
SITE 1622
SITE 1623
SITE 1624
SITE 1625
SITE 1626
SITE 1627
SITE 1628
SITE 1629
SITE 1630
SITE 1631
SITE 1632
SITE 1633
SITE 1634
SITE 1635
SITE 1636
SITE 1637
SITE 1638
SITE 1639
SITE 1640
SITE 1641
SITE 1642
SITE 1643
SITE 1644
SITE 1645
SITE 1646
SITE 1647
SITE 1648
SITE 1649
SITE 1650
SITE 1651
SITE 1652
SITE 1653
SITE 1654
SITE 1655
SITE 1656
SITE 1657
SITE 1658
SITE 1659
SITE 1660
SITE 1661
SITE 1662
SITE 1663
SITE 1664
SITE 1665
SITE 1666
SITE 1667
SITE 1668
SITE 1669
SITE 1670
SITE 1671
SITE 1672
SITE 1673
SITE 1674
SITE 1675
SITE 1676
SITE 1677
SITE 1678
SITE 1679
SITE 1680
SITE 1681
SITE 1682
SITE 1683
SITE 1684
SITE 1685
SITE 1686
SITE 1687
SITE 1688
SITE 1689
SITE 1690
SITE 1691
SITE 1692
SITE 1693
SITE 1694
SITE 1695
SITE 1696
SITE 1697
SITE 1698
SITE 1699
SITE 1700
SITE 1701
SITE 1702
SITE 1703
SITE 1704
SITE 1705
SITE 1706
SITE 1707
SITE 1708
SITE 1709
SITE 1710
SITE 1711
SITE 1712
SITE 1713
SITE 1714
SITE 1715
SITE 1716
SITE 1717
SITE 1718
SITE 1719
SITE 1720
SITE 1721
SITE 1722
SITE 1723
SITE 1724
SITE 1725
SITE 1726
SITE 1727
SITE 1728
SITE 1729
SITE 1730
SITE 1731
SITE 1732
SITE 1733
SITE 1734
SITE 1735
SITE 1736
SITE 1737
SITE 1738
SITE 1739
SITE 1740
SITE 1741
SITE 1742
SITE 1743
SITE 1744
SITE 1745
SITE 1746
SITE 1747
SITE 1748
SITE 1749
SITE 1750
SITE 1751
SITE 1752
SITE 1753
SITE 1754
SITE 1755
SITE 1756
SITE 1757
SITE 1758
SITE 1759
SITE 1760
SITE 1761
SITE 1762
SITE 1763
SITE 1764
SITE 1765
SITE 1766
SITE 1767
SITE 1768
SITE 1769
SITE 1770
SITE 1771
SITE 1772
SITE 1773
SITE 1774
SITE 1775
SITE 1776
SITE 1777
SITE 1778
SITE 1779
SITE 1780
SITE 1781
SITE 1782
SITE 1783
SITE 1784
SITE 1785
SITE 1786
SITE 1787
SITE 1788
SITE 1789
SITE 1790
SITE 1791
SITE 1792
SITE 1793
SITE 1794
SITE 1795
SITE 1796
SITE 1797
SITE 1798
SITE 1799
SITE 1800
SITE 1801
SITE 1802
SITE 1803
SITE 1804
SITE 1805
SITE 1806
SITE 1807
SITE 1808
SITE 1809
SITE 1810
SITE 1811
SITE 1812
SITE 1813
SITE 1814
SITE 1815
SITE 1816
SITE 1817
SITE 1818
SITE 1819
SITE 1820
SITE 1821
SITE 1822
SITE 1823
SITE 1824
SITE 1825
SITE 1826
SITE 1827
SITE 1828
SITE 1829
SITE 1830
SITE 1831
SITE 1832
SITE 1833
SITE 1834
SITE 1835
SITE 1836
SITE 1837
SITE 1838
SITE 1839
SITE 1840
SITE 1841
SITE 1842
SITE 1843
SITE 1844
SITE 1845
SITE 1846
SITE 1847
SITE 1848
SITE 1849
SITE 1850
SITE 1851
SITE 1852
SITE 1853
SITE 1854
SITE 1855
SITE 1856
SITE 1857
SITE 1858
SITE 1859
SITE 1860
SITE 1861
SITE 1862
SITE 1863
SITE 1864
SITE 1865
SITE 1866
SITE 1867
SITE 1868
SITE 1869
SITE 1870
SITE 1871
SITE 1872
SITE 1873
SITE 1874
SITE 1875
SITE 1876
SITE 1877
SITE 1878
SITE 1879
SITE 1880
SITE 1881
SITE 1882
SITE 1883
SITE 1884
SITE 1885
SITE 1886
SITE 1887
SITE 1888
SITE 1889
SITE 1890
SITE 1891
SITE 1892
SITE 1893
SITE 1894
SITE 1895
SITE 1896
SITE 1897
SITE 1898
SITE 1899
SITE 1900
SITE 1901
SITE 1902
SITE 1903
SITE 1904
SITE 1905
SITE 1906
SITE 1907
SITE 1908
SITE 1909
SITE 1910
SITE 1911
SITE 1912
SITE 1913
SITE 1914
SITE 1915
SITE 1916
SITE 1917
SITE 1918
SITE 1919
SITE 1920
SITE 1921
SITE 1922
SITE 1923
SITE 1924
SITE 1925
SITE 1926
SITE 1927
SITE 1928
SITE 1929
SITE 1930
SITE 1931
SITE 1932
SITE 1933
SITE 1934
SITE 1935
SITE 1936
SITE 1937
SITE 1938
SITE 1939
SITE 1940
SITE 1941
SITE 1942
SITE 1943
SITE 1944
SITE 1945
SITE 1946
SITE 1947
SITE 1948
SITE 1949
SITE 1950
SITE 1951
SITE 1952
SITE 1953
SITE 1954
SITE 1955
SITE 1956
SITE 1957
SITE 1958
SITE 1959
SITE 1960
SITE 1961
SITE 1962
SITE 1963
SITE 1964
SITE 1965
SITE 1966
SITE 1967
SITE 1968
SITE 1969
SITE 1970
SITE 1971
SITE 1972
SITE 1973
SITE 1974
SITE 1975
SITE 1976
SITE 1977
SITE 1978
SITE 1979
SITE 1980
SITE 1981
SITE 1982
SITE 1983
SITE 1984
SITE 1985
SITE 1986
SITE 1987
SITE 1988
SITE 1989
SITE 1990
SITE 1991
SITE 1992
SITE 1993
SITE 1994
SITE 1995
SITE 1996
SITE 1997
SITE 1998
SITE 1999
SITE 2000
SITE 2001
SITE 2002
SITE 2003
SITE 2004
SITE 2005
SITE 2006
SITE 2007
SITE 2008
SITE 2009
SITE 2010
SITE 2011
SITE 2012
SITE 2013
SITE 2014
SITE 2015
SITE 2016
SITE 2017
SITE 2018
SITE 2019
SITE 2020
SITE 2021
SITE 2022
SITE 2023
SITE 2024
SITE 2025
SITE 2026
SITE 2027
SITE 2028
SITE 2029
SITE 2030
SITE 2031
SITE 2032
SITE 2033
SITE 2034
SITE 2035
SITE 2036
SITE 2037
SITE 2038
SITE 2039
SITE 2040
SITE 2041
SITE 2042
SITE 2043
SITE 2044
SITE 2045
SITE 2046
SITE 2047
SITE 2048
SITE 2049
SITE 2050
SITE 2051
SITE 2052
SITE 2053
SITE 2054
SITE 2055
SITE 2056
SITE 2057
SITE 2058
SITE 2059
SITE 2060
SITE 2061
SITE 2062
SITE 2063
SITE 2064
SITE 2065
SITE 2066
SITE 2067
SITE 2068
SITE 2069
SITE 2070
SITE 2071
SITE 2072
SITE 2073
SITE 2074
SITE 2075
SITE 2076
SITE 2077
SITE 2078
SITE 2079
SITE 2080
SITE 2081
SITE 2082
SITE 2083
SITE 2084
SITE 2085
SITE 2086
SITE 2087
SITE 2088
SITE 2089
SITE 2090
SITE 2091
SITE 2092
SITE 2093
SITE 2094
SITE 2095
SITE 2096
SITE 2097
SITE 2098
SITE 2099
SITE 2100
SITE 2101
SITE 2102
SITE 2103
SITE 2104
SITE 2105
SITE 2106
SITE 2107
SITE 2108
SITE 2109
SITE 2110
SITE 2111
SITE 2112
SITE 2113
SITE 2114
SITE 2115
SITE 2116
SITE 2117
SITE 2118
SITE 2119
SITE 2120
SITE 2121
SITE 2122
SITE 2123
SITE 2124
SITE 2125
SITE 2126
SITE 2127
SITE 2128
SITE 2129
SITE 2130
SITE 2131
SITE 2132
SITE 2133
SITE 2134
SITE 2135
SITE 2136
SITE 2137
SITE 2138
SITE 2139
SITE 2140
SITE 2141
SITE 2142
SITE 2143
SITE 2144
SITE 2145
SITE 2146
SITE 2147
SITE 2148
SITE 2149
SITE 2150
SITE 2151
SITE 2152
SITE 2153
SITE 2154
SITE 2155
SITE 2156
SITE 2157
SITE 2158
SITE 2159
SITE 2160
SITE 2161
SITE 2162
SITE 2163
SITE 2164
SITE 2165
SITE 2166
SITE 2167
SITE 2168
SITE 2169
SITE 2170
SITE 2171
SITE 2172
SITE 2173
SITE 2174
SITE 2175
SITE 2176
SITE 2177
SITE 2178
SITE 2179
SITE 2180
SITE 2181
SITE 2182
SITE 2183
SITE 2184
SITE 2185
SITE 2186
SITE 2187
SITE 2188
SITE 2189
SITE 2190
SITE 2191
SITE 2192
SITE 2193
SITE 2194
SITE 2195
SITE 2196
SITE 2197
SITE 2198
SITE 2199
SITE 2200
SITE 2201
SITE 2202
SITE 2203
SITE 2204
SITE 2205
SITE 2206
SITE 2207
SITE 2208
SITE 2209
SITE 2210
SITE 2211
SITE 2212
SITE 2213
SITE 2214
SITE 2215
SITE 2216
SITE 2217
SITE 2218
SITE 2219
SITE 2220
SITE 2221
SITE 2222
SITE 2223
SITE 2224
SITE 2225
SITE 2226
SITE 2227
SITE 2228
SITE 2229
SITE 2230
SITE 2231
SITE 2232
SITE 2233
SITE 2234
SITE 2235
SITE 2236
SITE 2237
SITE 2238
SITE 2239
SITE 2240
SITE 2241
SITE 2242
SITE 2243
SITE 2244
SITE 2245
SITE 2246
SITE 2247
SITE 2248
SITE 2249
SITE 2250
SITE 2251
SITE 2252
SITE 2253
SITE 2254
SITE 2255
SITE 2256
SITE 2257
SITE 2258
SITE 2259
SITE 2260
SITE 2261
SITE 2262
SITE 2263
SITE 2264
SITE 2265
SITE 2266
SITE 2267
SITE 2268
SITE 2269
SITE 2270
SITE 2271
SITE 2272
SITE 2273
SITE 2274
SITE 2275
SITE 2276
SITE 2277
SITE 2278
SITE 2279
SITE 2280
SITE 2281
SITE 2282
SITE 2283
SITE 2284
SITE 2285
SITE 2286
SITE 2287
SITE 2288
SITE 2289
SITE 2290
SITE 2291
SITE 2292
SITE 2293
SITE 2294
SITE 2295
SITE 2296
SITE 2297
SITE 2298
SITE 2299
SITE 2300
SITE 2301
SITE 2302
SITE 2303
SITE 2304
SITE 2305
SITE 2306
SITE 2307
SITE 2308
SITE 2309
SITE 2310
SITE 2311
SITE 2312
SITE 2313
SITE 2314
SITE 2315
SITE 2316
SITE 2317
SITE 2318
SITE 2319
SITE 2320
SITE 2321
SITE 2322
SITE 2323
SITE 2324
SITE 2325
SITE 2326
SITE 2327
SITE 2328
SITE 2329
SITE 2330
SITE 2331
SITE 2332
SITE 2333
SITE 2334
SITE 2335
SITE 2336
SITE 2337
SITE 2338
SITE 2339
SITE 2340
SITE 2341
SITE 2342
SITE 2343
SITE 2344
SITE 2345
SITE 2346
SITE 2347
SITE 2348
SITE 2349
SITE 2350
SITE 2351
SITE 2352
SITE 2353
SITE 2354
SITE 2355
SITE 2356
SITE 2357
SITE 2358
SITE 2359
SITE 2360
SITE 2361
SITE 2362
SITE 2363
SITE 2364
SITE 2365
SITE 2366
SITE 2367
SITE 2368
SITE 2369
SITE 2370
SITE 2371
SITE 2372
SITE 2373
SITE 2374
SITE 2375
SITE 2376
SITE 2377
SITE 2378
SITE 2379
SITE 2380
SITE 2381
SITE 2382
SITE 2383
SITE 2384
SITE 2385
SITE 2386
SITE 2387
SITE 2388
SITE 2389
SITE 2390
SITE 2391
SITE 2392
SITE 2393
SITE 2394
SITE 2395
SITE 2396
SITE 2397
SITE 2398
SITE 2399
SITE 2400
SITE 2401
SITE 2402
SITE 2403
SITE 2404
SITE 2405
SITE 2406
SITE 2407
SITE 2408
SITE 2409
SITE 2410
SITE 2411
SITE 2412
SITE 2413
SITE 2414
SITE 2415
SITE 2416
SITE 2417
SITE 2418
SITE 2419
SITE 2420
SITE 2421
SITE 2422
SITE 2423
SITE 2424
SITE 2425
SITE 2426
SITE 2427
SITE 2428
SITE 2429
SITE 2430
SITE 2431
SITE 2432
SITE 2433
SITE 2434
SITE 2435
SITE 2436
SITE 2437
SITE 2438
SITE 2439
SITE 2440
SITE 2441
SITE 2442
SITE 2443
SITE 2444
SITE 2445
SITE 2446
SITE 2447
SITE 2448
SITE 2449
SITE 2450
SITE 2451
SITE 2452
SITE 2453
SITE 2454
SITE 2455
SITE 2456
SITE 2457
SITE 2458
SITE 2459
SITE 2460
SITE 2461
SITE 2462
SITE 2463
SITE 2464
SITE 2465
SITE 2466
SITE 2467
SITE 2468
SITE 2469
SITE 2470
SITE 2471
SITE 2472
SITE 2473
SITE 2474
SITE 2475
SITE 2476
SITE 2477
SITE 2478
SITE 2479
SITE 2480
SITE 2481
SITE 2482
SITE 2483
SITE 2484
SITE 2485
SITE 2486
SITE 2487
SITE 2488
SITE 2489
SITE 2490
SITE 2491
SITE 2492
SITE 2493
SITE 2494
SITE 2495
SITE 2496
SITE 2497
SITE 2498
SITE 2499
SITE 2500
SITE 2501
SITE 2502
SITE 2503
SITE 2504
SITE 2505
SITE 2506
SITE 2507
SITE 2508
SITE 2509
SITE 2510
SITE 2511
SITE 2512
SITE 2513
SITE 2514
SITE 2515
SITE 2516
SITE 2517
SITE 2518
SITE 2519
SITE 2520
SITE 2521
SITE 2522
SITE 2523
SITE 2524
SITE 2525
SITE 2526
SITE 2527
SITE 2528
SITE 2529
SITE 2530
SITE 2531
SITE 2532
SITE 2533
SITE 2534
SITE 2535
SITE 2536
SITE 2537
SITE 2538
SITE 2539
SITE 2540
SITE 2541
SITE 2542
SITE 2543
SITE 2544
SITE 2545
SITE 2546
SITE 2547
SITE 2548
SITE 2549
SITE 2550
SITE 2551
SITE 2552
SITE 2553
SITE 2554
SITE 2555
SITE 2556
SITE 2557
SITE 2558
SITE 2559
SITE 2560
SITE 2561
SITE 2562
SITE 2563
SITE 2564
SITE 2565
SITE 2566
SITE 2567
SITE 2568
SITE 2569
SITE 2570
SITE 2571
SITE 2572
SITE 2573
SITE 2574
SITE 2575
SITE 2576
SITE 2577
SITE 2578
SITE 2579
SITE 2580
SITE 2581
SITE 2582
SITE 2583
SITE 2584
SITE 2585
SITE 2586
SITE 2587
SITE 2588
SITE 2589
SITE 2590
SITE 2591
SITE 2592
SITE 2593
SITE 2594
SITE 2595
SITE 2596
SITE 2597
SITE 2598
SITE 2599
SITE 2600
SITE 2601
SITE 2602
SITE 2603
SITE 2604
SITE 2605
SITE 2606
SITE 2607
SITE 2608
SITE 2609
SITE 2610
SITE 2611
SITE 2612
SITE 2613
SITE 2614
SITE 2615
SITE 2616
SITE 2617
SITE 2618
SITE 2619
SITE 2620
SITE 2621
SITE 2622
SITE 2623
SITE 2624
SITE 2625
SITE 2626
SITE 2627
SITE 2628
SITE 2629
SITE 2630
SITE 2631
SITE 2632
SITE 2633
SITE 2634
SITE 2635
SITE 2636
SITE 2637
SITE 2638
SITE 2639
SITE 2640
SITE 2641
SITE 2642
SITE 2643
SITE 2644
SITE 2645
SITE 2646
SITE 2647
SITE 2648
SITE 2649
SITE 2650
SITE 2651
SITE 2652
SITE 2653
SITE 2654
SITE 2655
SITE 2656
SITE 2657
SITE 2658
SITE 2659
SITE 2660
SITE 2661
SITE 2662
SITE 2663
SITE 2664
SITE 2665
SITE 2666
SITE 2667
SITE 2668
SITE 2669
SITE 2670
SITE 2671
SITE 2672
SITE 2673
SITE 2674
SITE 2675
SITE 2676
SITE 2677
SITE 2678
SITE 2679
SITE 2680
SITE 2681
SITE 2682
SITE 2683
SITE 2684
SITE 2685
SITE 2686
SITE 2687
SITE 2688
SITE 2689
SITE 2690
SITE 2691
SITE 2692
SITE 2693
SITE 2694
SITE 2695
SITE 2696
SITE 2697
SITE 2698
SITE 2699
SITE 2700
SITE 2701
SITE 2702
SITE 2703
SITE 2704
SITE 2705
SITE 2706
SITE 2707
SITE 2708
SITE 2709
SITE 2710
SITE 2711
SITE 2712
SITE 2713
SITE 2714
SITE 2715
SITE 2716
SITE 2717
SITE 2718
SITE 2719
SITE 2720
SITE 2721
SITE 2722
SITE 2723
SITE 2724
SITE 2725
SITE 2726
SITE 2727
SITE 2728
SITE 2729
SITE 2730
SITE 2731
SITE 2732
SITE 2733
SITE 2734
SITE 2735
SITE 2736
SITE 2737
SITE 2738
SITE 2739
SITE 2740
SITE 2741
SITE 2742
SITE 2743
SITE 2744
SITE 2745
SITE 2746
SITE 2747
SITE 2748
SITE 2749
SITE 2750
SITE 2751
SITE 2752
SITE 2753
SITE 2754
SITE 2755
SITE 2756
SITE 2757
SITE 2758
SITE 2759
SITE 2760
SITE 2761
SITE 2762
SITE 2763
SITE 2764
SITE 2765
SITE 2766
SITE 2767
SITE 2768
SITE 2769
SITE 2770
SITE 2771
SITE 2772
SITE 2773
SITE 2774
SITE 2775
SITE 2776
SITE 2777
SITE 2778
SITE 2779
SITE 2780
SITE 2781
SITE 2782
SITE 2783
SITE 2784
SITE 2785
SITE 2786
SITE 2787
SITE 2788
SITE 2789
SITE 2790
SITE 2791
SITE 2792
SITE 2793
SITE 2794
SITE 2795
SITE 2796
SITE 2797
SITE 2798
SITE 2799
SITE 2800
SITE 2801
SITE 2802
SITE 2803
SITE 2804
SITE 2805
SITE 2806
SITE 2807
SITE 2808
SITE 2809
SITE 2810
SITE 2811
SITE 2812
SITE 2813
SITE 2814
SITE 2815
SITE 2816
SITE 2

Query Match 100.0%; Score 66; DB 1; Length 1527;
 Best Local Similarity 100.0%; Pred. No. 0.00038;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHVALNTP 13
 DB 1350 QPVLHVALNTP 1362

RESULT 2
 CALH_HUMAN
 ID CALH_HUMAN STANDARD; PRT; 1516 AA.
 AC P39060; Q9UK38; Q9Y6Q7; Q9Y6Q8;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 16-OCT-2001 (Rel. 46, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Collagen alpha 1(XVIII) chain precursor [Contains: Endostatin].
 GN COL18A1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98164096; PubMed=9503365;
 RA Saarela J., Ylikarppa R., Rehn M., Purmonen S., Pihlajaniemi T.;
 RT "Complete primary structure of two variant forms of human type XVIII
 collagen and tissue-specific differences in the expression of the
 corresponding transcripts.";
 RL Matrix Biol. 16:319-328(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20289799; PubMed=10830953;
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
 RA Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
 RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
 RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordiek G., Hornischer K., Brandt P.,
 RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
 RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
 RA Wehmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
 RA Lehrach H., Reinhardt R., Yaspo M.-L.;
 RT "The DNA sequence of human chromosome 21.";
 RL Nature 405:311-319(2000).
 RN [3]
 RP SEQUENCE OF 834-1516 FROM N.A.
 RX MEDLINE=94245237; PubMed=8188291;
 RA Oh S.P., Warman M.L., Seldin M.F., Cheng S., Knoll J.H., Timmons S.,
 RA Olsen B.R.;
 RT "Cloning of cDNA and genomic DNA encoding human type XVIII collagen
 and localization of the alpha 1(XVIII) collagen gene to mouse
 chromosome 10 and human chromosome 21.";
 RL Genomics 19:494-499(1994).
 RN [4]
 RP SEQUENCE OF 1334-1516 FROM N.A.
 RC TISSUE=Placenta;
 RA Zhi-Yong H., Biao L., Wei-Jie Z., Xiang-Fu W.;
 RT "Cloning and expression of human endostatin gene in Escherichia
 coli.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP INVOLVEMENT IN KNOBLOCH SYNDROME.
 RX MEDLINE=20400145; PubMed=10942434;
 RA Sertie A.L., Sossi V., Camargo A.A., Zatz M., Brahe C.,
 RA Passos-Bueno M.R.;
 RT "Collagen XVIII, containing an endogenous inhibitor of angiogenesis
 and tumor growth, plays a critical role in the maintenance of retinal
 structure and in neural tube closure.";
 RL Hum. Mol. Genet. 9:2051-2058(2000).

[6]
 VARIANT ASN-1437.
 RX MEDLINE=21518361; PubMed=11606364;
 RA Iughetti P., Suzuki O., Godoi P.H., Alves V.A., Sertie A.L.,
 RA Zorick T., Soares F., Camargo A., Moreira E.S., di Loreto C.,
 RA Moreira-Filho C.A., Simpson A., Oliva G., Passos-Bueno M.R.;
 RT "A polymorphism in endostatin, an angiogenesis inhibitor, predisposes
 to the development of prostatic adenocarcinoma.";
 RL Cancer Res. 61:7375-7378(2001).
 CC -!- FUNCTION: COL18A1 PROBABLY PLAYS A MAJOR ROLE IN DETERMINING THE
 RETINAL STRUCTURE AS WELL AS IN THE CLOSURE OF THE NEURAL TUBE.
 CC -!- FUNCTION: ENDOSTATIN POTENTIALLY INHIBITS ENDOTHELIAL CELL
 PROLIFERATION AND ANGIOGENESIS. MAY INHIBIT ANGIOGENESIS BY
 BINDING TO THE HEPARAN SULPHATE PROTEOGLYCAN INVOLVED IN GROWTH
 FACTOR SIGNALLING.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long; Synonyms=NC-493;
 CC IsoId=P39060-1; Sequence=Displayed;
 CC Name=Short; Synonyms=NC1-303;
 CC IsoId=P39060-2; Sequence=VSP_001155, VSP_001156;
 CC -!- TISSUE SPECIFICITY: PRESENT IN MULTIPLE ORGANS WITH HIGHEST LEVELS
 IN LIVER, LUNG AND KIDNEY.
 CC -!- PTM: Prolines at the third position of the tripeptide repeating
 unit (G-X-Y) are hydroxylated in some or all of the chains.
 CC -!- POLYMORPHISM: There is an association between a polymorphism in
 position 1437 and prostate cancer. Heterozygous Asn-1437
 individuals have a 2.5 times increased chance of developing
 prostate cancer as compared with homozygous Asp-1437 individuals.
 CC -!- DISEASE: Defects in COL18A1 are a cause of Knobloch syndrome (KNO).
 occurrence of high myopia, vitreoretinal degeneration with retinal
 detachment, macular abnormalities and occipital encephalocele.
 CC -!- SIMILARITY: BELONGS TO THE FIBRIL-ASSOCIATED COLLAGENS WITH
 INTERRUPTED HELICES (FACIT) FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF018081; AAC39658.1; -;
 DR EMBL; AF018082; AAC39659.1; -;
 DR EMBL; AL163302; CAB90482.1; -;
 DR EMBL; L22548; AAA51864.1; -;
 DR EMBL; AF184060; AAF01310.1; ALT_INIT.
 DR PUB; IENL; 02-DEC-98.
 DR GlycoSuiteDB; P39060;
 DR Genew; HGNC:2195; COL18A1.
 DR MIM; 120328; -;
 DR MIM; 267750; -;
 DR GO; GO:0005581; C:collagen; TAS.
 DR GO; GO:0008181; F:tumor suppressor; TAS.
 DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.
 DR GO; GO:0008245; P:negative regulation of cell proliferation; TAS.
 DR GO; GO:0007048; P:oncogenesis; TAS.
 DR GO; GO:0007601; P:vision; TAS.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR001791; Laminin_G.
 DR InterPro; IPR03129; TSPN.
 DR Pfam; PF01391; Collagen; 7.
 DR Pfam; PF02210; TSPN; 1.
 DR ProDom; PDC00007; Clg_helix; 1.
 DR SMART; SM00282; LamG; 1.
 DR SMART; SM00210; TSPN; 1.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Cell adhesion; Collagen; Glycoprotein; Signal; Alternative splicing;
 KW Polymorphism; 3D-structure. POTENTIAL.
 FT SIGNAL 1 23
 FT CHAIN 24 1516 COLLAGEN ALPHA 1 (XVIII) CHAIN.

CHAIN	1334	1516	Query Match
DOMAIN	24	516	93.9%;
DOMAIN	517	550	Best Local Similarity 92.3%;
DOMAIN	551	560	Pred. No. 0.0021;
DOMAIN	561	640	Mismatches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
DOMAIN	641	664	
DOMAIN	665	786	
DOMAIN	787	809	
DOMAIN	810	892	
DOMAIN	893	906	
DOMAIN	907	948	
DOMAIN	949	961	
DOMAIN	962	1034	
DOMAIN	1035	1044	
DOMAIN	1045	1077	
DOMAIN	1078	1089	
DOMAIN	1090	1111	
DOMAIN	1112	1118	
DOMAIN	1119	1173	
DOMAIN	1174	1186	
DOMAIN	1187	1204	
DOMAIN	1205	1516	
CARBOHYD	68	68	
CARBOHYD	129	129	
CARBOHYD	164	164	
CARBOHYD	691	691	
CARBOHYD	1329	1329	
DISULFID	1366	1506	
DISULFID	1468	1498	
SITE	1095	1097	
VARSPLIC	1	180	
VARSPLIC	181	215	
VARIANT	1437	1437	
CONFLICT	428	428	
CONFLICT	841	841	
CONFLICT	877	877	
CONFLICT	886	886	
CONFLICT	912	912	
CONFLICT	933	933	
CONFLICT	975	975	
CONFLICT	1064	1064	
CONFLICT	1084	1084	
CONFLICT	1120	1120	
CONFLICT	1123	1123	
CONFLICT	1126	1126	
CONFLICT	1206	1206	
CONFLICT	1304	1304	
CONFLICT	1314	1314	
CONFLICT	1323	1324	
CONFLICT	1443	1443	
CONFLICT	1483	1483	
SEQUENCE	1516 AA;	153840 MW;	
Query Match			Score 62; DB 1; Length 1516;
Best Local Similarity			92.3%;
Mismatches 12; Conservative			1; Mismatches 0; Indels 0; Gaps 0;
QY	1 QPVLHLVALNTEPL 13		
Db	1340 QPVLHLVALNTEPL 1352		
RESULT 3			
CALE_HUMAN			
ID_CALE_HUMAN			
STANDARD;			
1388 AA.			
PRT;			

FT DOMAIN 733 763 NONHELICAL REGION 3 (NC3).
 FT DOMAIN 754 798 TRIPLE-HELICAL REGION 3 (COL3).
 FT DOMAIN 799 822 TRIPLE-HELICAL REGION 4 (NC4).
 FT DOMAIN 823 867 TRIPLE-HELICAL REGION 4 (COL4).
 FT DOMAIN 868 878 TRIPLE-HELICAL REGION 5 (NC5).
 FT DOMAIN 879 949 TRIPLE-HELICAL REGION 5 (COL5).
 FT DOMAIN 950 983 TRIPLE-HELICAL REGION 6 (NC6).
 FT DOMAIN 984 1013 TRIPLE-HELICAL REGION 6 (COL6).
 FT DOMAIN 1014 1027 TRIPLE-HELICAL REGION 7 (NC7).
 FT DOMAIN 1028 1045 TRIPLE-HELICAL REGION 7 (COL7).
 FT DOMAIN 1046 1052 TRIPLE-HELICAL REGION 8 (NC8).
 FT DOMAIN 1053 1107 TRIPLE-HELICAL REGION 8 (COL8).
 FT DOMAIN 1108 1117 TRIPLE-HELICAL REGION 9 (NC9).
 FT DOMAIN 1118 1132 TRIPLE-HELICAL REGION 9 (COL9).
 FT DOMAIN 1133 1398 NONHELICAL REGION 10 (NC10).
 FT REPEAT 358 555 4 X TANDEM REPEATS.
 FT REPEAT 358 408 1.
 FT REPEAT 409 459 2.
 FT REPEAT 460 509 3.
 FT REPEAT 510 555 4.
 FT CARBOHYD 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 324 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 687 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 807 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 814 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1046 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 10 10 C -> S (IN REF. 2).
 FT CONFLICT 49 49 D -> V (IN REF. 2).
 FT CONFLICT 95 95 L -> A (IN REF. 2).
 FT CONFLICT 150 150 P -> A (IN REF. 2).
 FT CONFLICT 204 204 M -> V (IN REF. 2).
 FT CONFLICT 409 409 R -> A (IN REF. 2).
 SQ SEQUENCE 1388 AA; 141930 MW; 60822AD925A3093D CRC64;

 Query Match 66.7%; Score 44; DB 1; Length 1388;
 Best Local Similarity 66.7%; Pred. No. 4.4;
 Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

 QY 1 QPVLHLVALNTP 12
 DB 1215 KPALHLAALNMP 1226

 RESULT 4
 MOT1 YEAST
 ID MOT1 YEAST STANDARD; PRT; 1867 AA.
 AC P32333.
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable helicase MOT1.
 GN MOT1 OR YPL082C OR LIF4C.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92195335; PubMed=1312673;
 RA Davis J.L., Kunisawa R., Thorne J.;
 RT "A presumptive helicase (MOT1 gene product) affects gene expression
 and is required for viability in the yeast Saccharomycetes
 cerevisiae.";
 RT Mol. Cell. Biol. 12:1879-1892(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AB972;
 RX MEDLINE=97313271; PubMed=9169875;
 RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansong W.,
 RA Araujo R., Aparicio A., Barrell B.G., Badcock K., Benes V.,
 RA Botstein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
 RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
 RA Dietrich F.S., Delius H., Dipaolo T., Dubois E., Duesterhoeft A.,

RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
 RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.,
 RA Hunnicke-Smith S., Hyman R., Johnston M., Kalman S., Klein K.,
 RA Komp C., Kurdi C., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
 RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
 RA Mueller-Auer S., Namath A., Newtich U., Oefner P., Pearson D.,
 RA Petel F.X., Pohl T.M., Purnelle D., Schafer M., Scharfe M.,
 RA Scherens B., Schramm S., Schroeder M., Sidu A.M., Tettelin H.,
 RA Urescarazu L.A., Ushinsky S., Vierdeels F., Vissers S., Voss H.,
 RA Walsh S.V., Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E.,
 RA Zhong W.W., Zollner A., Vo D.H., Hani J.;
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI";
 RL Nature 387:103-105(1997).
 CC -1- FUNCTION: REGULATES TRANSCRIPTION IN ASSOCIATION WITH TATA BINDING
 CC PROTEIN (TBP). REMOVES TBP FROM THE TATA BOX IN AN ATP-DEPENDENT
 CC MANNER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: BELONGS TO THE SNF2/RAD54 HELICASE FAMILY.
 CC -1- SIMILARITY: Contains 6 HEAT repeats.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M83224; AAA34786.1; -;
 DR EMBL; U41849; AAB86257.1; -;
 DR PIR; S22775; S22775;
 DR TRAFSPAC; TQ3499; -;
 DR SGD; S0006003; MOT1.
 DR GO; GO:0000228; C:nuclear chromosome; IDA.
 DR GO; GO:0016887; F:ATPase activity; IDA.
 DR GO; GO:0006357; P:regulation of transcription from Pol II pro. .; IDA.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR000357; HEAT repeat.
 DR InterPro; IPR001650; Helicase C.
 DR InterPro; IPR000330; SNF2 N_
 DR Pfam; PF00271; helicase C_1.
 DR Pfam; PF00176; SNF2 N; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELIC; 1.
 DR PROSITE; PS00077; HEAT REPEAT; FALSE NEG.
 DR KW Nuclear protein; DNA-binding; Helicase; ATP-binding; Repeat.
 FT DOMAIN 195 211 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT REPEAT 289 326 HEAT 1.
 FT REPEAT 445 482 HEAT 2.
 FT REPEAT 541 578 HEAT 3.
 FT REPEAT 1108 1145 HEAT 4.
 FT REPEAT 1188 1225 HEAT 5.
 FT REPEAT 1495 1537 HEAT 6.
 FT NP BIND 1297 1304 ATP (POTENTIAL).
 FT SITE 1408 1411 DEGH BOX.
 SQ SEQUENCE 1867 AA; 209975 MW; 1A00005148D5632B CRC64;

 Query Match 65.2%; Score 43; DB 1; Length 1867;
 Best Local Similarity 69.2%; Pred. No. 9.4;
 Matches 9; Conservative 2; Mismatches 0; Indels 2; Gaps 1;

 QY 1 QPVLHLVALNTP 13
 DB 637 QPILHL--LNTPV 647

 RESULT 5
 ATP6 RHOPD
 ID ATP6 RHOPD STANDARD; PRT; 217 AA.
 AC Q986H6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

DE ATP synthase A chain (EC 3.6.3.14) (Protein 6).
GN MTATP6 OR ATP6.
OS Rhopalosiphum padi (Bird cherry-oat aphid).
OG Mitochondrion.
CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
CC Aphidoidea; Aphididae; Aphidini; Rhopalosiphum.
OX NCBI_TaxID=40932;
RN [1]
RP SEQUENCE FROM N.A.
RA Martinez-Torres D., Buades C., Latorre A., Moya A.;
RT "Molecular systematics of aphids and their primary endosymbionts.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Key component of the proton channel; it may play a
CC direct role in the translocation of protons across the membrane
CC (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out).
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(1) has five
CC subunits: alpha(3), beta(3), gamma(1), delta(1), epsilon(1). CF(0)
CC has three main subunits: A, B and C (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ298673; CAC28069.1;
DR InterPro: IPR000568; ATPSynt_Asub.
DR Pfam: PF00119; ATP-synt A; 1.
DR PRINTS: PR00123; ATPASEA.
DR TIGRFS: TIGR01131; ATP synt_6_or_A; 1.
DR PROSITE: PS00449; ATPASEA; 1. Mitochondrion; Transmembrane.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
SQ SEQUENCE 217 AA; 25442 MW; 2F3EEEAAC53B89B4 CRC64;

Query Match 50.6%; Score 40; DB 1; Length 217;
Best Local Similarity 54.5%; Pred. No. 3.3;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 VLHLVALNTP1 13
DB 127 IAHLIPLNTP1 137

RESULT 6
3MGH-DEIRA STANDARD; PRT; 190 AA.
AC Q9RSQ0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative 3-methyladenine DNA glycosylase (EC 3.2.2.-).
DR D2074.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=2036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

```

```

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
CC -!- SIMILARITY: Belongs to the DNA glycosylase MPG family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE002043; AAF11623.1;
DR PIR: D75319; D75319.
DR HSP; P29372; IBNK.
DR TIGR; DR2074; -.
DR HAMAP; MF_09527; -.
DR InterPro: IPR003180; PurDNA_glycosylase.
DR Pfam: PF02245; PurDNA_glyco; 1.
DR ProDom: PD009649; PurDNA_glycosylase; 1.
DR Hypothetical protein; DNA repair; Hydrolase; Complete proteome.
KW Hypothetical protein; DNA repair; Hydrolase; Complete proteome.
SQ SEQUENCE 190 AA; 20819 MW; 05264B48A76B6B2 CRC64;

Query Match 59.1%; Score 39; DB 1; Length 190;
Best Local Similarity 66.7%; Pred. No. 4.4;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PVLHLVALNTP1 13
DB 136 PEHLJAPETPL 147

RESULT 7
POLG-PSBMV STANDARD; PRT; 3206 AA.
AC P29152;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Genome polyprotein [Contains: N-terminal protein (PI); Helper
DE component proteinase (EC 3.4.22.45) (HC-Pro); Protein P3; 6 kDa
DE protein 1 (6k1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2
DE (6k2); Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A)
DE (NIA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-Pro); Nuclear
DE inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase)
DE (EC 2.7.7.48); Coat protein (CP)].
OS pea seed-borne mosaic virus (strain DPD1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
OC Potyvirus.
OX NCBI_TaxID=31736;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9204431; PubMed=1940858;
RA Johansen B., Rasmussen O.F., Heide M., Borkhardt B.;
RT "The complete nucleotide sequence of pea seed-borne mosaic virus
RT RNA.";
RL J. Gen. Virol. 72:2625-2632(1991).
CC -!- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC -!- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
CC MAY BE INVOLVED IN REPLICATION.
CC -!- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC -!- CATALYTIC ACTIVITY: Hydrolyzes glutamyl bonds, and activity is
CC further restricted by preferences for the amino acids in P6 - P1,
CC that vary with the species of potyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-
CC Gin+(Ser or Gly) for the enzyme from tobacco etch virus. The
CC natural substrate is the viral polyprotein, but other proteins and
CC oligopeptides containing the appropriate consensus sequence are
CC also cleaved.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

```

[RNA] (N).
-!- CATALYTIC ACTIVITY: Hydrolyzes a Gly-|-Gly bond at its own C-terminus, commonly in the sequence -Tyr-Xaa-Val-Gly-|-Gly, in the processing of the potyviral polyprotein.
-!- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
-!- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT INDIVIDUAL PROTEINS.
-!- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
-!- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
-!- SIMILARITY: BELONGS TO THE POTYVIRUS POLYPEPTIDE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D10930; BAA01726.1; -
PIR; JQ1331; GNVSEF.
MEROPS; C04.010; -
MEROPS; C06.001; -
InterPro; IPR001410; DEAD.
InterPro; IPR001450; Helicase_C.
InterPro; IPR001720; Peptidase_C4.
InterPro; IPR001456; Peptidase_C6.
InterPro; IPR001592; Poty_coat.
InterPro; IPR002540; Poty_P1.
InterPro; IPR007095; RNA_pol_PS.
InterPro; IPR001205; RNA_pol_P3d.
InterPro; IPR007094; RNA_pol_Psvir.
Pfam; PF00271; helicase_C; 1.
Pfam; PF00863; Peptidase_C4; 1.
Pfam; PF00851; Peptidase_C6; 1.
Pfam; PF00767; Poty_coat; 1.
Pfam; PF01577; Poty_P1; 1.
Pfam; PF00680; RNA_dep_RNA_pol; 1.
PRINTS; PR00966; NIAPOTYPTASE.
SMART; SM00489; DEXD; 1.
SMART; SM00490; HELIC_C; 1.
Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
KW ATP-binding.
FT CHAIN 1 ? N-TERMINAL PROTEIN.
FT CHAIN 2 ? HELPER COMPONENT PROTEINASE.
FT CHAIN 857 ? PROTEIN P3
FT CHAIN ? 1266 6 kDa PROTEIN 1.
FT CHAIN 1267 1902 CYTOPLASMIC INCLUSION PROTEIN.
FT CHAIN 1903 1955 6 kDa PROTEIN 2.
FT CHAIN 1956 ? GENOME-LINKED PROTEIN.
FT CHAIN ? 2395 ? NUCLEAR INCLUSION PROTEIN A.
FT CHAIN 2396 2915 NUCLEAR INCLUSION PROTEIN B.
FT CHAIN 2916 3206 COAT PROTEIN.
FT CHAIN 1266 1267 CLEAVAGE (BY 49 kDa PROTEASE).
FT SITE 1902 1903 CLEAVAGE (BY 49 kDa PROTEASE).
FT SITE 1955 1956 CLEAVAGE (BY 49 kDa PROTEASE).
FT SITE 2395 2396 CLEAVAGE (BY 49 kDa PROTEASE).
FT SITE 2915 2916 CLEAVAGE (BY 49 kDa PROTEASE).
FT BINDING 2016 2016 COVALENT LINKAGE OF VIRAL RNA (BY SIMILARITY).
FT NP BIND 1351 1358 ATP (POTENTIAL).
SQ SEQUENCE 3206 AA; 364271 MW; 42A3D921BEE9A0CBF CRC64;

Query Match 57.6%; Score 38; DB 1; Length 3206;
Best Local Similarity 50.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 PVLHVALNTP 13
DB 67 PINHVDKSTPI 78

RESULT 8
YM21_YEAST STANDARD; PRT; 102 AA.
ID YN21_YEAST
AC F40211;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical 12.1 kDa protein in MDS1-RPL13B intergenic region.
DE YMR141C OR YMR375.10C.
GN Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T., Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S., Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A., Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RA "The nucleotide sequence of Saccharomyces cerevisiae chromosome XIII.";
RT Nature 387:90-93 (1997).
RL Nature 387:90-93 (1997).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Z47071; CAA87355.1; -
PIR; S50397; S50397.
SGD; S0004749; YMR141C.
KW Hypothetical protein.
SQ SEQUENCE 102 AA; 12108 MW; 8A7F6449CC86F219 CRC64;

Query Match 56.1%; Score 37; DB 1; Length 102;
Best Local Similarity 50.0%; Pred. No. 5.3;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 VHLHVALNTP 12
DB 29 IMHLISNTP 38

```

RX MEDLINE=87284169; PubMed=3039177;
RA Pfitzner A.J., Tsai E.C., Strominger J.L., Speck S.H.;
RT "Isolation and characterization of cDNA clones corresponding to
RT transcripts from the BamHI H and F regions of the Epstein-Barr virus
RL genome.";
RL J. Virol. 61:2902-2909 (1987).
RN [3]
RP IDENTIFICATION OF PROTEIN.
RA MEDLINE=87321098; PubMed=2820125;
RA Pearson G.R., Luka J., Petti L., Sample J., Birkenbach M., Braun D.,
RA Kieff E.;
RT "Identification of an Epstein-Barr virus early gene encoding a second
RT component of the restricted early antigen complex.";
RL Virol. 160:151-161 (1987).
CC -!- FUNCTION: INHIBITS APOPTOSIS.
CC -!- MISCELLANEOUS: EA-R IS PART OF THE RESTRICTED EA-COMPLEX.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 1 (BH1) domain.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 2 (BH2) domain.
CC -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17416; AAA45873.1; -
CC DR EMBL; V01555; -; NOT ANNOTATED_CDS.
CC DR EMBL; M17293; AAA45875.1; -
CC DR EMBL; A22899; CAA01638.1; -
CC DR PIR; C93085; QQB84.
CC DR InterPro; IPR000712; Bcl2_BH.
CC DR InterPro; IPR002475; BCL2_family.
CC DR PROSITE; PS50062; BCL2_FAMILY; 1.
CC DR PROSITE; PS01080; BH1; FALSE_NEG.
CC DR PROSITE; PS01258; BH2; 1.
CC DR Early protein; Transmembrane; Apoptosis.
CC KW DOMAIN 89 109 BH1.
CC FT DOMAIN 142 157 BH2.
CC FT TRANSMEM 166 186 POTENTIAL.
CC FT CARBOHYD 22 22 N-LINKED (GLCNAC. .) (POTENTIAL).
CC FT CARBOHYD 118 118 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SQ SEQUENCE 191 AA; 21893 MW; 8108BCB94F81DC8B CRC64;
CC
CC Query Match 56.1%; Score 37; DB 1; Length 191;
CC Best Local Similarity 66.7%; Pred. No. 10;
CC Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
CC
CC QY 2 PVLHLVALNTP 13
CC Db 27 PVLELAARETPL 38
CC
CC RESULT 10
CC SKI MOUSE STANDARD; PRT; 348 AA.
CC ID SKI_MOUSE
CC AC Q60598;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Ski oncogene (C-ski) (Fragment).
CC GN SKI.
CC OS Mus musculus (Mouse).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC OX NCBI_TaxID=10090;
CC [1]
CC RN SEQUENCE FROM N.A.
CC RP STRAIN=129/J;
CC RC MEDLINE=96127473; PubMed=8573720;
CC RX Namciu S., Lyons G.E., Micales B.K., Heyman H.-C., Colmenares C.,
RA

```

```

RA Stavnezer E.;
RT "Enhanced expression of mouse c-ski accompanies terminal skeletal
RT muscle differentiation in vivo and in vitro.";
RL Dev. Dyn. 204:291-300 (1995).
CC -!- FUNCTION: May play a role in terminal differentiation of skeletal
CC muscle cells but not in the determination of cells to the myogenic
CC lineage.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DEVELOPMENTAL STAGE: Is expressed in a uniform pattern in all
CC embryonic cells prior to skeletal muscle cell formation in the
CC myotomes of somites. Expression is first upregulated in skeletal
CC muscle at 12 dpc, this upregulation is evident first in body wall
CC muscle and one day later in limb muscles. At 13.5 dpc a most
CC prominent expression is seen in all skeletal muscles. At this
CC stage expression is seen in all other cells and tissues but at
CC lower levels than in skeletal muscle.
CC -!- SIMILARITY: TO SKI PROTEIN OF SLOAN-KETTERING VIRUS AND TO SNO
CC ONCOGENES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U14173; AAA39669.1; -
CC DR MGD; MGI:98310; Ski.
CC DR GO; GO:005737; C:cytoplasm; IDA.
CC DR GO; GO:0005634; C:nucleus; IDA.
CC DR GO; GO:0017053; C:transcriptional repressor complex; IDA.
CC DR GO; GO:0005515; F:protein binding activity; IPI.
CC DR GO; GO:0030326; P:limb morphogenesis; IMP.
CC DR InterPro; IPR003380; Transform_Ski.
CC DR Pfam; PF02437; Ski_Sno; 1.
CC DR Proto-oncogene; Nuclear protein; Repeat.
CC KW NON_TER 348 348
CC FT SEQUENCE 348 AA; 37869 MW; 4DD73478145D038C CRC64;
CC SQ
CC
CC Query Match 56.1%; Score 37; DB 1; Length 348;
CC Best Local Similarity 63.6%; Pred. No. 20;
CC Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
CC
CC QY 2 PVLHLVALNTP 12
CC Db 70 PVLLHPLAIOFP 80
CC
CC RESULT 11
CC SKI_HUMAN STANDARD; PRT; 728 AA.
CC ID SKI_HUMAN
CC AC P12755;
CC DT 01-OCT-1989 (Rel. 12, Created)
CC DT 01-OCT-1989 (Rel. 12, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE Ski oncogene (C-ski).
CC GN SKI.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC [1]
CC RN SEQUENCE FROM N.A.
CC RP MEDLINE=93345144; PubMed=2762147;
CC RA Nomura N., Sasamoto S., Ishii S., Matsui M., Ishizaki R.;
CC RT "Isolation of human cDNA clones of ski and the ski-related gene,
CC RT sno.";
CC RL Nucleic Acids Res. 17:5489-5500 (1989).
CC -!- FUNCTION: May play a role in terminal differentiation of skeletal
CC muscle cells but not in the determination of cells to the myogenic
CC lineage.
CC -!- SUBUNIT: INTERACTS WITH SMAD2, 3 AND 4.
CC

```


CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: TO SKI PROTEIN OF SLOAN-KETTERING VIRUS AND TO SNO
CC ONCOGENES
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X15218, CAA33268.1; -
CC PIR: S06053; TVRUSK.
CC PDB: 1MR1; 21-JAN-03.
CC TRANSFAC: T04643; -
CC Genew: HGNC:10896; SKI.
CC MIM: 164780; -
CC DR GO: GO:0005634; C:nucleus; NAS
CC DR GO: GO:0030154; P:cell differentiation; NAS.
CC DR InterPro: IPR003360; TransForm_Ski.
CC DR Pfam: PF02437; Ski_Sno; 1.
CC KW Proto-oncogene; Coiled coil; Nuclear protein; Repeat; 3D-structure.
CC FT DOMAIN 536 710 COILED COIL (POTENTIAL).
CC SQ SEQUENCE 728 AA; 80004 MW; 9578C484DA28C2DA CRC64;
Query Match 56.1%; Score 37; DB 1; Length 728;
Best Local Similarity 63.6%; Pred. No. 45;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 PVLHLVALNTP 12
DB 72 PVLHLPAOPP 82
RESULT 12
VF13_BPH6
ID VP13_BPH6 STANDARD; PRT; 72 AA.
AC P11130;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1999 (Rel. 11, Last sequence update)
DT 01-JAN-1990 (Rel. 13, Last annotation update)
DE P13 protein.
GN P13.
OS Bacteriophage phi-6.
OC Viruses; dsRNA viruses; Cystoviridae; Cystovirus.
OX NCBI_TaxID=10879;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-10.
RX MEDLINE=8160044; PubMed=3247997;
RA Gottlieb P., Metzger S., Romantschuk M., Carton J., Strassman J.,
RA Bamford D.H., Kalkinen N., Mindich L.;
RT "Nucleotide sequence of the middle dsRNA segment of bacteriophage phi
RT 6: placement of the genes of membrane-associated proteins.";
RL Virology 163:183-190(1988).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M17462; AAA68486.1; -
CC PIR: D28648; PNBPF6.
CC KW Envelope protein.
SQ SEQUENCE 72 AA; 7649 MW; 61DEAB3B71053B88 CRC64;
Query Match 54.5%; Score 36; DB 1; Length 72;
Best Local Similarity 58.3%; Pred. No. 5.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTP 12
DB 14 QPLVLVALNTP 25
RESULT 13
UCP4_HUMAN
ID UCP4_HUMAN STANDARD; PRT; 323 AA.
AC O95647;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitochondrial uncoupling protein 4 (UCP 4).
GN UCP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RC MEDLINE=99148824; PubMed=10025957;
RA Mao W., Yu X.X., Zhong A., Li W., Brush J., Sherwood S.W., Adams S.H.,
RA Pan G.;
RT "UCP4, a novel brain-specific mitochondrial protein that reduces
RT membrane potential in mammalian cells.";
RL FEBS Lett 443:326-330(1999).
CC -!- FUNCTION: UCP ARE MITOCHONDRIAL TRANSPORTER PROTEINS THAT CREATE
CC PROTON LEAKS ACROSS THE INNER MITOCHONDRIAL MEMBRANE, THUS
CC UNCOUPLING OXIDATIVE PHOSPHORYLATION FROM ATP SYNTHESIS. AS A
CC RESULT, ENERGY IS DISSIPATED IN THE FORM OF HEAT. MAY PLAY A ROLE
CC IN THERMOREGULATORY HEAT PRODUCTION AND METABOLISM IN BRAIN.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
CC inner membrane.
CC -!- TISSUE SPECIFICITY: FOUND IN ADULT AND FETAL BRAIN. PRESENT IN
CC MOST OF THE BRAIN TISSUES, WITH LOW LEVELS IN SPINAL CHORD, CORPUS
CC CALLOSUM AND SUBSTANTIA NIGRA.
CC -!- SIMILARITY: Belongs to the mitochondrial carrier family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF110532; AAD16995.1; -
CC DR GO: GO:0005739; C:mitochondrion; TAS.
CC DR GO: GO:0015302; F:uncoupling protein activity; TAS.
CC DR GO: GO:0006091; P:energy pathways; TAS.
CC DR InterPro: IPR002030; Mit_uncoupling.
CC DR InterPro: IPR001993; Mitoch_carrier.
CC DR Pfam: PF00153; mito_carr; 3.
CC DR PRINTS: PR00784; MTUNCOUPLING.
CC DR PROSITE: PS00215; MITOCH_CARRIER; 2.
KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
FT TRANSMEM 23 40 POTENTIAL.
FT TRANSMEM 88 109 POTENTIAL.
FT TRANSMEM 127 144 POTENTIAL.
FT TRANSMEM 195 212 POTENTIAL.
FT TRANSMEM 229 248 POTENTIAL.
FT TRANSMEM 288 311 POTENTIAL.
FT DOMAIN 5 8 POLY-GLU.
SQ SEQUENCE 323 AA; 36064 MW; 4C54A56BB10333ED CRC64;
Query Match 54.5%; Score 36; DB 1; Length 323;
Best Local Similarity 63.6%; Pred. No. 28;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 3 VHLHLVALNTP 13
DB 214 VKHLVLNTP 224

```
Search completed: January 26, 2004, 19:09:31
Job time : 27 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 26, 2004, 19:04:49 ; Search time 83 Seconds

(without alignments)
40.418 Million cell updates/sec

Title: US-09-766-412-30

Perfect score: 66
Sequence: 1 QPVLHLVALNTPL 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_page:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_invertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	66	100.0	184	11 Q9JK63	Q9JK63 mus musculus
2	66	100.0	1140	11 Q61434	Q61434 mus musculus
3	66	100.0	1774	11 Q82001	Q82001 mus musculus
4	62	93.9	187	4 Q8W415	Q8W415 homo sapien
5	62	93.9	261	4 Q8NG19	Q8NG19 homo sapien
6	62	93.9	816	4 Q8N4S4	Q8N4S4 homo sapien
7	62	93.9	1344	13 Q93419	Q93419 gallus gall
8	61	92.4	226	11 Q3Q2D2	Q3Q2D2 rattus norv
9	59	89.4	1367	11 Q3EOD9	Q3EOD9 mus musculus
10	59	89.4	1367	11 Q35206	Q35206 mus musculus
11	54	81.8	171	11 Q9WUW5	Q9WUW5 rattus norv
12	52	78.8	1315	13 Q8OHL9	Q8OHL9 xenopus lae
13	50	75.8	1307	13 Q8JFF7	Q8JFF7 xenopus lae
14	45	68.2	217	8 Q9B6C6	Q9B6C6 eriosoma la
15	45	68.2	361	13 Q8AMC6	Q8AMC6 brachydanio
16	45	68.2	1388	4 Q9Y4W4	Q9Y4W4 homo sapien

17	43	65.2	217	8 Q9B6P8	Q9B6P8 geioica utri
18	43	65.2	218	8 Q9B6G9	Q9B6G9 tetraneura
19	41	62.1	568	11 Q9Z1N7	Q9Z1N7 mus musculu
20	40	60.6	176	16 O05584	O05584 mycobacteri
21	40	60.6	217	8 Q9B7T1	Q9B7T1 apionura 1
22	40	60.6	217	8 Q9B612	Q9B612 panaphis ju
23	40	60.6	217	8 Q9B7C6	Q9B7C6 acyrthosiph
24	40	60.6	217	8 Q9B6Q8	Q9B6Q8 drepanosiph
25	40	60.6	217	8 Q9B6H4	Q9B6H4 schizolachn
26	40	60.6	217	8 Q9B6H5	Q9B6H5 schizaphis
27	40	60.6	217	8 Q9B6K0	Q9B6K0 lachnus rob
28	40	60.6	217	8 Q9B6G3	Q9B6G3 thelaxus su
29	40	60.6	217	8 Q9B6Q1	Q9B6Q1 eulachnus s
30	40	60.6	217	8 Q9B615	Q9B615 pemphigus b
31	40	60.6	217	8 Q9B6H9	Q9B6H9 phylloxera
32	40	60.6	217	8 Q9B610	Q9B610 pemphigus s
33	40	60.6	217	8 Q9B6Q0	Q9B6Q0 forda formi
34	40	60.6	217	8 Q9B6H8	Q9B6H8 phylloxera
35	40	60.6	217	8 Q9B6R3	Q9B6R3 chaetophoru
36	40	60.6	217	8 Q9B6J3	Q9B6J3 myzus persi
37	40	60.6	217	8 Q9B6G4	Q9B6G4 tuberolachn
38	40	60.6	342	16 Q8ZKB8	Q8ZKB8 salmonella
39	40	60.6	342	16 Q8Z1A2	Q8Z1A2 salmonella
40	40	60.6	1538	10 Q94H26	Q94H26 cryza sativ
41	39	59.1	268	17 Q8TM56	Q8TM56 methanosarc
42	39	59.1	333	2 Q46143	Q46143 clostridium
43	39	59.1	435	16 Q8Z530	Q8Z530 salmonella
44	39	59.1	494	4 Q96HR8	Q96HR8 homo sapien
45	39	59.1	533	10 Q9LP55	Q9LP55 arabidopsis

ALIGNMENTS

RESULT 1

Q9JK63 ID Q9JK63 PRELIMINARY; PRT; 184 AA.
AC Q9JK63; 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Endostatin (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chinese Kunming;
RA Jia S., Zhu F., Xing G., Yu Y., Duan C., Xiu R.-J., He F.;
RT "Anticancer treatment of targeted fusion protein delivery to tumor
RT neovasculation."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF257775; AAF69009.1; -.
DR HSSP; P39061; 1KOE.
FT NON_TER 1 184
SQ SEQUENCE 184 AA; 20376 MW; AC06F9D8D103412A CRC64;
Query Match 100.0%; Score 66; DB 11; Length 184;
Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13

Db 7 QPVLHLVALNTPL 19

RESULT 2

Q61434 ID Q61434 PRELIMINARY; PRT; 1140 AA.
AC Q61434;
DT 01-NOV-1996 (TREMBlrel. 01, Created)

```

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Collagen (Fragment).
GN COL15A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Abe N., Muragaki Y., Yoshioka H., Inoue H., Ninomiya Y.;
RT "Identification of a novel collagen chain represented by extensive
RT interruptions in the triple-helical region.";
RL Cell. Mol. Biol. Res. 196:576-582(1993).
DR EMBL; D15462; BAA04483.1; -.
DR HSSP; P39061; 1KOE.
DR MGD; MGI:88451; Coll18a1.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01391; Collagen; 8.
DR Pfam; PF02210; TSPN; 1.
DR ProDom; PD000007; Collagen; 1.
DR Collagen.
KW Collagen.
FT NON TER
SQ SEQUENCE 1140 AA; 115156 MW; 8B0C7E6862B3BDFE CRC64;

Query Match 100.0%; Score 66; DB 11; Length 1140;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 963 QPVLHLVALNTPL 975

RESULT 3
Q62001 PRELIMINARY; PRT; 1774 AA.
AC Q62001; Q60672;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Procollagen, type XVIII, alpha 1 precursor (XVIII) collagen
DE (Procollagen, type XVIII, alpha 1) (Alpha-1 type XVIII collagen).
GN COL18A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Rehn M., Hintikka E., Pihlajaniemi T.;
RT "Primary structure of the alpha 1 chain of mouse type XVIII collagen,
RT partial structure of the corresponding gene, and comparison of the
RT alpha 1 (XVIII) chain with its homologue, the alpha 1(XV) collagen
RT chain.";
RL J. Biol. Chem. 269:13929-13935(1994).
RN [2]
RP SEQUENCE OF 1-562 FROM N.A.
RX MEDLINE=94240112; PubMed=8189894;
RA Rehn M., Pihlajaniemi T.;
RT "Alpha 1 (XVIII), a collagen chain with frequent interruptions in the
RT collagenous sequence, a distinct tissue distribution, and homology
RT with type XV collagen.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:4234-4238(1994).
RN [3]
RP SEQUENCE OF 1-562 FROM N.A.
RX MEDLINE=95181468; PubMed=7876242;
RA Rehn M., Pihlajaniemi T.;
RT "Identification of three N-terminal ends of type XVIII collagen chains
RT and tissue-specific differences in the expression of the corresponding

```

```

RT transcripts. The longest form contains a novel motif homologous to rat
RT and Drosophila frizzled proteins";
RL J. Biol. Chem. 270:4705-4711(1995).
DR EMBL; U03715; AAC52903.1; -.
DR EMBL; U03716; AAC52903.1; JOINED.
DR EMBL; U03718; AAC52903.1; JOINED.
DR EMBL; U34607; AAC52903.1; JOINED.
DR EMBL; U34608; AAC52903.1; JOINED.
DR EMBL; U34609; AAC52903.1; JOINED.
DR EMBL; U34610; AAC52903.1; JOINED.
DR EMBL; U34611; AAC52903.1; JOINED.
DR EMBL; U34612; AAC52903.1; JOINED.
DR EMBL; U34613; AAC52903.1; JOINED.
DR EMBL; U11637; AAC52179.1; -.
DR HSSP; P39061; 1KOE.
DR MGD; MGI:88451; Coll18a1.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR000087; Fz domain.
DR InterPro; IPR001791; Laminin-G.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01391; Collagen; 8.
DR Pfam; PF01392; Fz; 1.
DR Pfam; PF02210; TSPN; 1.
DR ProDom; PD000007; Collagen; 1.
DR SMART; SM00063; FRI; 1.
DR SMART; SM00282; LamG; 1.
DR SMART; SM00210; TSPN; 1.
DR PROSITE; PS00038; FZ; 1.
DR Collagen; Signal.
KW Collagen.
SQ SEQUENCE 1774 AA; 182229 MW; CF4D9BC9E88EF232 CRC64;

Query Match 100.0%; Score 66; DB 11; Length 1774;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 1597 QPVLHLVALNTPL 1609

RESULT 4
Q8WXI5 PRELIMINARY; PRT; 187 AA.
AC Q8WXI5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Collagen XVII (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21409408; PubMed=11517600;
RA Feng Y., Cui L.B., Liu C.X., Ma Q.J.;
RT "Inhibition effect in vitro of purified endostatin expressed in Pichia
RT pastoris.";
RL Sheng Wu Gong Cheng Xue Bao 17:278-282(2001).
DR EMBL; AF416592; AAL37720.1; -.
FT NON TER
SQ SEQUENCE 187 AA; 20448 MW; 72B1047D8938CD3 CRC64;

Query Match 93.9%; Score 62; DB 4; Length 187;
Best Local Similarity 92.3%; Pred. No. 0.0014;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
DB 11 QPVLHLVALNTPL 23

RESULT 5

```

Q8NG19 Q8NG19 PRELIMINARY; PRT; 261 AA.
 AC Q8NG19;
 DT 01-OCT-2002 (T-EMBLrel. 22, Created)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Multi-functional protein MFP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dou D.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF282883; AAMS2249.1; -
 SQ SEQUENCE 261 AA; 26745 MW; CA60C920AF3E90E5 CRC64;

Query Match 93.9%; Score 62; DB 4; Length 261;
 Best Local Similarity 92.3%; Pred. No. 0.002;
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
 Db 85 QPVLHLVALNSPL 97
 |||||

RESULT 6
 Q8N4S4 Q8N4S4 PRELIMINARY; PRT; 816 AA.
 AC Q8N4S4;
 DT 01-OCT-2002 (T-EMBLrel. 22, Created)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Similar to collagen, type XVIII, alpha 1 (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC033715; AAH33715.1; -
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF01391; Collagen; 6.
 DR ProDom; PD000007; Collagen; 1.
 KW Collagen.
 FT NON_TER
 SQ SEQUENCE 816 AA; 82553 MW; 5D53982946694F86 CRC64;

Query Match 93.9%; Score 62; DB 4; Length 816;
 Best Local Similarity 92.3%; Pred. No. 0.0063;
 Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
 Db 640 QPVLHLVALNSPL 652
 |||||

RESULT 7
 O93419 O93419 PRELIMINARY; PRT; 1344 AA.
 AC O93419;
 DT 01-NOV-1998 (T-EMBLrel. 08, Created)
 DT 01-OCT-2001 (T-EMBLrel. 18, Last sequence update)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
 DE Collagen XVIII precursor.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.

OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98411346; PubMed=9738008;
 RA Halfter W., Dong S., Schurer B., Cole G.J.;
 RT "Collagen XVIII is a basement membrane heparan sulfate proteoglycan."
 RL J. Biol. Chem. 273:25404-25412(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Halfter W., Dong S.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF083440; AAC33294.2; -
 DR HSSP; P39061; IKOE.
 DR InterPro; IPR000087; Collagen.
 DR InterPro; IPR001791; Laminin G.
 DR InterPro; IPR003129; TSPN.
 DR Pfam; PF01391; Collagen; 8.
 DR Pfam; PF02210; TSPN; 1.
 DR ProDom; PD000007; Collagen; 2.
 DR SMART; SM00282; LamG; 1.
 DR SMART; SM00210; TSPN; 1.
 KW Collagen; Signal.
 FT SIGNAL
 SQ SEQUENCE 1344 AA; 137402 MW; 7AA366B4FE940CCD CRC64;

Query Match 93.9%; Score 62; DB 13; Length 1344;
 Best Local Similarity 92.3%; Pred. No. 0.01;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 QPVLHLVALNTPL 13
 Db 1167 QPALHLVALNTPL 1179
 |||||

RESULT 8
 Q9QZD2 Q9QZD2 PRELIMINARY; PRT; 226 AA.
 AC Q9QZD2;
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
 DE Collagen XVIII (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=20227226; PubMed=10766159;
 RA Perletti G., Concarl P., Giardini R., Marras E., Piccinini F.,
 RA Folkman J., Chen L.;
 RT "Antitumor activity of endostatin against carcinogen-induced rat
 RT primary mammary tumors."
 RL Cancer Res. 60:11793-11796(2000).
 DR EMBL; AF189709; AAF00975.1; -
 DR HSSP; P39061; IKOE.
 FT NON_TER
 SQ SEQUENCE 226 AA; 25350 MW; 38B83C0486C0E949 CRC64;

Query Match 92.4%; Score 61; DB 11; Length 226;
 Best Local Similarity 100.0%; Pred. No. 0.0026;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PVLHLVALNTPL 13
 Db 50 PVLHLVALNTPL 61
 |||||

RESULT 9
 Q9EQD9 Q9EQD9 PRELIMINARY; PRT; 1367 AA.
 ID Q9EQD9
 AC Q9EQD9;

```

RESULT 10
O35206 PRELIMINARY; PRT; 1367 AA.
AC O35206;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Type XV collagen.
GN COL15A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
GX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97480713; PubMed=9339358;
RA Hagg P.W., Korelli-Kuitunen N., Eklund L., Palotie A.,
RA Pihlajaniemi T.;
RT Cloning of mouse type XV collagen sequences and mapping of the
RT corresponding gene to 4p1-3. Comparison of mouse and human alpha 1
RT (XV) collagen sequences indicates divergence in the number of small
RT collagenous domains.";
RL Genomics 45:31-41(1997).
DR ENBL; AF011450; AAC53387.1; -.
DR HSSP; P39061; 1KOE
DR MGD; MGI:88449; Collisal.
DR InterPro; IPR000087; Collagen.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR003129; TSPN.
DR Pfam; PF01391; Collagen; 5.
DR Pfam; PF02210; TSPN; 1.
DR ProDom; PD000007; Collagen; 1.
DR SMART; SM00282; LamG; 1.
DR SMART; SM00210; TSPN; 1.
DR COLLAGEN.
KW KW
SQ SEQUENCE 1367 AA; 140525 MW; A483A1254AF3ABEC CRC64;

Query Match 89.4%; Score 59; DB 11; Length 1367;
Best Local Similarity 84.6%; Pred. No. 0.039;
Matches 11; Conservative 2; Mismatches 0; Indels 0; Gaps

Qy 1 QPVLHVLVAINTPL 13
Db 1194 RPYLHVLVAINTPV 1206

RESULT 11
Q9WUW5 PRELIMINARY; PRT; 171 AA.
AC Q9WUW5
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Collagen type XVIIi, alpha (I) chain (Fragment).
GN COL18A1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CX NCBI_TaxID=10116;
GX [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RA Jia J.D., Bauer M., Eberspaecher U., Donner P., Schuppan D.;
RT "Partial 3'-terminal cDNA sequence of rat collagen XVIIi/endostatin."
RL Submitted (FEB-1999) to the ENBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Liver;
RA Jia J.D., Bauer M., Sedlaczek N., Ruehl M., Riecken E.O., Schuppan D.;
RT "Temporopspatial expression of collagen XVIIi/endostatin in acute and
RT chronic liver injuries."
RL Submitted (FEB-1999) to the ENBL/GenBank/DBJ databases.
DR ENBL; A02316873; CAB44263.1;

```

```

DR HSSP: P39061; 1KOE.
DR NON_TER 1
FT NON_TER 171
SQ SEQUENCE 171 AA; 18933 MW; 81BE2EB3FC2C8E72 CRC64;

Query Match      81.8%; Score 54; DB 11; Length 171;
Best Local Similarity 100.0%; Pred. No. 0.04;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VLHLVALNTPL 13
   |||||
Db 1 VLHLVALNTPL 11

RESULT 12
QBQHL9 PRELIMINARY; PRT; 1315 AA.
AC QBQHL9
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Type XVII collagen alpha1 chain.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Ishino T., Sekimizu K., Natori S., Kubo T.;
RT "Identification and characterization of genes expressed selectively in
RT the regenerating tail of Xenopus laevis tadpole.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB047066; BAB84674.1; -.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR003129; TSPN.
DR Pfam: PF03391; Collagen; 7.
DR Pfam: PF02210; TSPN; 1.
DR ProDom: PD000007; Collagen; 1.
DR SMART; SM00210; TSPN; 1.
KW Collagen.
SQ SEQUENCE 1315 AA; 134946 MW; 0C56C235DE058365 CRC64;

Query Match      78.8%; Score 52; DB 13; Length 1315;
Best Local Similarity 83.3%; Pred. No. 0.75;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PVLHLVALNTPL 13
   |||||
Db 1139 PALHLVALNAPL 1150

RESULT 13
QBUFF7 PRELIMINARY; PRT; 1307 AA.
AC QBUFF7
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Type XVII collagen short variant.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22166979; PubMed=12175494;
RA Blamaa H., Peterson J., Pihlajaniemi T., Destree O.;
RT "Cloning of three variants of type XVII collagen and their expression
RT patterns during Xenopus laevis development.";
RL Mech. Dev. 114:109-113(2002).
DR EMBL; AY052763; AAL14257.1; -.

```

```

DR InterPro: IPR000087; Collagen.
DR InterPro: IPR003129; TSPN.
DR Pfam: PF01391; Collagen; 6.
DR Pfam: PF02210; TSPN; 1.
DR SMART; SM00210; TSPN; 1.
KW Collagen.
SQ SEQUENCE 1307 AA; 134153 MW; D53EDBFE3DE34976 CRC64;

Query Match      75.8%; Score 50; DB 13; Length 1307;
Best Local Similarity 75.0%; Pred. No. 1.8;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PVLHLVALNTPL 13
   |||||
Db 1131 PALHLVALNAPI 1142

RESULT 14
QB6Q6 PRELIMINARY; PRT; 217 AA.
AC QB6Q6
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE ATP synthase A chain subunit 6 (EC 3.6.3.14).
OS Eriosoma lanuginosum.
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Paraneoptera; Hemiptera; Sternorrhyncha; Aphidiformes;
OC Aphidoidea; Pemphigidae; Eriosomatinae; Eriosoma.
OX NCBI_TaxID=135953;
RN [1]
RP SEQUENCE FROM N.A.
RA Martinez-Torres D., Buades C., Latorre A., Moya A.;
RT "Molecular systematics of aphids and their primary endosymbionts.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A
CC DIRECT ROLE IN THE TRANSLOCATION OF PROTONS ACROSS THE MEMBRANE
CC (BY SIMILARITY).
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
DR EMBL; AJ298665; CAC27833.1; -.
DR InterPro: IPR000568; ATPsynth_Asub.
DR Pfam: PF00119; ATP-synt_A; 1.
DR PRINTS; PR0123; ATPASEA.
DR TIGRfams; TIGR01131; ATP_synt_6_or_A; 1.
DR CF(0); Hydrogen ion transport; Ion transport; Transmembrane;
KW Transport; Mitochondrion.
SQ SEQUENCE 217 AA; 25552 MW; DFFB88F5D5AD49C8 CRC64;

Query Match      68.2%; Score 45; DB 8; Length 217;
Best Local Similarity 63.6%; Pred. No. 2.5;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VLHLVALNTPL 13
   |||||
Db 127 IHLPLNTPL 137

RESULT 15
QB6WC6 PRELIMINARY; PRT; 361 AA.
AC QB6WC6
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Collagen XVIII (Fragment).
GN COL18A1.

```

OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Hattak Z., Morvan-Dubois G., Thisse B., Garrone R., Le Guellec D.;
RT "Sequence and embryonic expression of collagen XVIII NC11 domain
RL (endostatin) in the zebrafish.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ494837; CAD38825.1; -.
FT NON TER 1
SQ SEQUENCE 361 AA; 40222 MW; 3C5A0F8479D26735 CRC64;

Query Match 68.2%; Score 45; DB 13; Length 361;
Best Local Similarity 72.7%; Pred. No. 4.1;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PVLHLVALNTP 12
| | | | | | | | | |
Db 185 PGLHLIALNSP 195

Search completed: January 26, 2004, 15:08:03
Job time : 88 secs